

複数ユーザで動的実環境を共有するための移動ロボットを用いた遠隔体験システム

著者	曹 暢
内容記述	筑波大学修士（情報学）学位論文・平成28年3月25日授与（36021号）
発行年	2016
学位授与年度	2015
URL	http://hdl.handle.net/2241/00145440

複数ユーザで動的実環境を共有するための
移動ロボットを用いた遠隔体験システム

筑波大学大学院
図書館情報メディア研究科
2016 年 3 月
曹 暢

目次

第1章	序論	2
1.1	研究の背景	2
1.2	遠隔体験サービス	2
1.3	研究の目的	3
1.4	本論文の構成	3
第2章	関連研究と本研究の提案	4
2.1	遠隔体験サービス	4
2.1.1	道路周辺画像提供サービス	4
2.1.2	遠隔操作移動ロボット	6
2.1.3	ライブカメラ	7
2.1.4	現存サービスの問題点	8
2.2	遠隔コミュニケーション	8
2.2.1	ノンバーバルコミュニケーション	9
2.2.2	環境情報	9
2.3	自由視点生成	11
2.3.1	モデルベースレンダリング法(MBR)	11
2.3.2	イメージベースレンダリング法(IBR)	12
2.4	本研究の位置付け	14
第3章	遠隔共同体験システムの構成	15
3.1	システム概要	15
3.2	システム構成	16
3.3	ハードウェア構成	17
3.3.1	移動ロボット	19
3.3.2	カメラ	19
3.3.3	LRF	20
3.3.4	Kinect for Windows	22
3.3.5	ゲームパッド	23
3.4	ソフトウェア構成	24
第4章	共同体験バーチャル空間生成のアルゴリズム	25
4.1	共同体験バーチャル空間の構成	25
4.2	共同体験バーチャル空間生成の基本的な流れ	26
4.3	位置固定環境の生成	27
4.3.1	位置固定環境生成の流れ	27
4.3.2	SLAMを用いた環境地図と自己位置推定	28
4.3.3	位置固定環境のモデルの生成	29
4.3.4	テクスチャマッピング	30
4.4	歩行者の描画	35

4.5	同行者の描画	37
第5章	実験	38
5.1	実験環境	38
5.2	実験の条件	40
5.2.1	システム 1: 共同体験バーチャル空間を更新しない遠隔体験	41
5.2.2	システム 2: アバタを利用しない遠隔体験	42
5.2.3	システム 3: 自由視点を生成しない遠隔体験	42
5.2.4	システム 4: 実環境の情報を基づき共同体験バーチャル空間を更新する遠隔体験	43
5.3	実験の手順	43
5.3.1	実験 1	43
5.3.2	実験 2 (追加実験)	45
5.4	実験の評価方法	48
5.5	実験の結果	49
5.5.1	アバタの効果検証	49
5.5.2	環境情報更新の効果検証	51
5.5.3	自由視点の効果検証	53
第6章	結論	56
	謝辞	57
	参考文献	58
	付録	60

図目次

図 2.1	グーグルストリームビュー	4
図 2.2	長野県のバーチャルリアリティ観光体験	5
図 2.3	遠隔操作ロボットによる美術館巡りプロジェクト「After Dark」	6
図 2.4	世界中のライブ映像を生中継し続けるウェブサイト「EarthCam」	7
図 2.5	遠隔コミュニケーションにおける空間の使われ方	10
図 2.6	復元した 3 次元形状を基に生成した自由視点画像例 ([20] より)	11
図 2.7	モーフィング([21] より)	12
図 2.8	ライトフィールドレンダリングで生成した自由視点画像例 ([23] より)	13
図 3.1	共同遠隔体験システムのイメージ図	15
図 3.2	遠隔共同体験システムの構成	16
図 3.3	ロボットの構成	18
図 3.4	ユーザ側の構成	18
図 3.5	カメラの配置	20
図 3.6	LRF のスキャン範囲	21
図 3.7	Kinect for Windows v2	22
図 3.8	Kinect SDK2.0 の骨格データ	22
図 3.9	Gamepad の操作	23
図 3.10	ロボット側がユーザに送信する情報	24
図 4.1	遠隔体験システムのインターフェース一例	25
図 4.2	インターフェース生成の流れ	26
図 4.3	位置固定環境生成の流れ	27
図 4.4	ICP-SLAM により環境地図を作成	28
図 4.5	位置固定環境のモデルの生成結果	29
図 4.6	カメラのピンホールモデル	30
図 4.7	カメラ画像の歪み修正	32
図 4.8	カメラ画像の選択	34
図 4.9	位置固定環境の自由視点生成結果	34
図 4.10	LRF による人物追跡	36
図 4.11	歩行者モデルの生成結果	36
図 4.12	同行者アバタの構成	37
図 5.1	実験環境の俯瞰図	38
図 5.2	ユーザ操作風景	40
図 5.3	自由視点あり場合のミニマップ	41
図 5.4	システム 3 のインターフェース例	42
図 5.5	実験タスクの色紙の初期配置	45
図 5.6	追加実験のターゲットの初期配置	47
図 5.7	アバタ効果検証実験のアンケートの結果	51
図 5.8	システム 1 とシステム 4 のアンケートの結果	53
図 5.9	システム 3 とシステム 4 のアンケートの結果	55

表目次

表 3.1	移動ロボットと搭載するセンサ	17
表 3.2	使用されたカメラの仕様	19
表 3.3	使用される LRF の仕様	21
表 3.4	Kinect for Windows v2 のスペック	23
表 5.1	動作環境	39
表 5.2	使用したライブラリ	39
表 5.3	実験条件の設定	41
表 5.4	タスクの達成時間	49
表 5.5	タスクの達成時間	53

第1章 序論

1.1 研究の背景

近年，インターネットやパソコン，スマートフォン，タブレットなどの普及により，「いつでも」「どこでも」「誰とでも」つながれるネットワーク社会が実現しつつある．これらのインフラや情報端末などを利用し，遠隔地の情報を提供する様々なサービスが出現しつつある．中でも遠隔地の情報を提供し，家にいながらその地を歩くような感覚をユーザに与える遠隔体験サービスにおいて，様々なシステムが開発され，時間や金銭的な原因で今まで行けなかった場所でも，いつでもどこでも簡単に行ったような感覚を味わうことができるようになりつつある．このようなサービスはユーザに利益があるだけではなく，観光地側も，遠隔旅行サービスを通して地域情報を積極的に情報発信することができ，地域活性化につなげることが可能となる．また，これまで関心を持たなかった人たちの興味も引き出すことができ，新たな観光客の獲得につなげる可能性も考えられる．

1.2 遠隔体験サービス

自宅にいながら，名所やミュージアム、遊園地、お店などの遠隔地を楽しむサービスやイベントはすでにいくつが存在する．中でも Google Inc.が提供している「ストリートビュー」[1]は世界中の多くの地域情報を公開し，世の中に広く知れ渡ったサービスとなっている．他にも楽天(株)の旅行予約サイト「楽天トラベル」が提供する「バーチャルリアリティ（VR）観光体験」イベント，THE WORKERS (LONDON) LTD.が提供している「After Dark」[2]は，遠隔操作できる移動ロボットを通して，夜の美術館を巡るプロジェクトである．世界中に設置されたライブカメラを通して，その場所の「今」の映像をリアルタイムに見ることができるサービスなども多数ある．しかしこれらのサービスは旅行感覚を与えるにはそれぞれ不十分な点があり，またどのサービスも複数人が同行して一緒に旅行を楽しむ場合を想定していない．

1.3 研究の目的

本研究の目的は，ネットワークでつながった複数のユーザがお互いにコミュニケーションしつつ，実際に観光地に居るように，それぞれ自由に遠隔地を移動することができ，そして遠隔地の「今」の状況をリアルタイムに把握することが可能な，共同遠隔体験システムを研究開発することである．

コミュニケーションについては，単に音声やビデオ電話のような空間と無関係のものではなく，ユーザたちが同じ場（遠隔地）にいるような感覚がすることを目指す．また，遠隔地は少ない機器を用いてシステムの実現を目指す．

1.4 本論文の構成

本論文は本章を含め 6 章で構成される．各章の概要は次の通りである．第 1 章では，研究の背景と本研究が目指す目的について述べた．第 2 章では，関連研究を紹介する．第 3 章では，本研究で提案したアプローチのシステム概要と構成について述べる．第 4 章では，システムのインターフェースを生成するためのアルゴリズムについて論述する．第 5 章では，提案システムを評価する実験の内容と結果の考察について述べる．第 6 章では，実験の結果を踏まえ，まとめを述べる．

第2章 関連研究と本研究の提案

2.1 遠隔体験サービス

遠隔地の情報を提供し、現場の雰囲気を味わえるサービスはたくさん存在する。実現方法からは主に3つに分類することができる。以下はこれらについてそれぞれ紹介する。

2.1.1 道路周辺画像提供サービス

道路周辺画像提供サービスはグーグルストリートビューを代表に、マップのナビゲーションやバーチャルツアーなど、いろんな用途に使われるシステムである。他にもロケーションビュー、ウォークスルービデオシステム、Microsoft Street Slide などいろんな種類がある。

グーグルストリートビューのインターフェイス例を図 2.1 に示す。グーグルストリートビューとは Google Inc.が提供しているグーグルマップ及び Google Earth 上で利用できるサービスの1つであり、カメラから取得した複数の画像から重複部分を削除しつなぎ合わせることで1つの360度パノラマ画像として表示したものである。当初はアメリカの主要都市のみ対応していたが、現在はパリ、ロンドン、ニューヨーク、香港、東京といった世界各国の主要都市をカバーしている。ストリートビューで使用された画像はカメラを搭載した撮影車を用いて撮影されたものである。車内には、GPS 装置、速度計、方位計などのセンサが搭載されており、地図の正確な位置に画像を表示することができる。



図 2.1 グーグルストリートビュー

グーグルストリートビューと似たようなシステムで、バーチャル旅行に狙いを絞り開発されたシステムも既に存在する。

楽天株式会社の運営する旅行予約サイト「楽天トラベル」は、長野県と連携し、銀座 NAGANO にて楽天の研究開発部門である「楽天技術研究所」独自の技術を活用した「バーチャルリアリティ（VR）観光体験」イベントを、国内旅行会社として初めて開催した。

イベントの参加者たちは、アメリカで開発された「ヘッドマウントディスプレイ“Oculus Rift”」[3]を頭に装着し、視線を向ける方向に合わせて観光地の画像を提示するという仕掛けがある。画像は 360 度全方位に対応していて、Google のストリートビューを操作して観光地を巡るよりも、その場に実際にいて観光しているような自然な体験ができると思われる。



図 2.2 長野県のバーチャルリアリティ観光体験

2.1.2 遠隔操作移動ロボット

遠隔操作ロボットは、これまで主に宇宙探索や震災地のレスキューなどに利用されていたが、近年の IT 技術の発展に伴い、個人向きの製品も数多く製品化されるようになってきた[4]。観光方面に使われるロボットは案内ロボットのほか、ユーザの代わりに遠隔地を巡り、比較的到高自由度で場の映像をユーザに提供するものがある。

「IK Prize 2014」で 1 位になったプロジェクト「After Dark」は、遠隔操作ロボットに搭載されたカメラを通して、ロンドンにある有名な美術館「テート・ブリテン」に夜ひとりで過ごす疑似体験ができるようにした。カメラを搭載した 4 台のロボットが、夜の美術館を移動し、実況付きの動画ストリームを見ることができる。使用されたロボットとその遠隔操作インターフェースを図 2.3 に示す。移動ロボットを遠隔操作することにより、閉館後の美術館内を自由に見学することができるシステムである。



図 2.3 遠隔操作ロボットによる美術館巡りプロジェクト「After Dark」

2.1.3 ライブカメラ

現在世界の各地にライブカメラが設置されており，インターネットを通しライブカメラが写した世界の「今」を見ることができる．このサービスはさらに動画像と静止画像による提供に分かれ，動画像では，その場所の「今」を動いている映像として見ることができ，静止画像を提供するタイプのライブカメラのほとんどは，ユーザがアクセスした瞬間の風景写真が見れるようにしているものである．世界中のライブカメラが撮影する映像を，24 時間 365 日，生中継し続けるウェブサイトが「EarthCam」である[5]．



図 2.4 世界中のライブ映像を生中継し続けるウェブサイト「EarthCam」

2.1.4 現存サービスの問題点

以上のシステムを利用すれば遠隔体験が実現できるが、直接現地を訪ねる観光と比べると、事前に撮影した画像を利用するストリートビューや楽天トラベルのバーチャルリアリティ観光体験システムは過去の状態しか提供できず、現地のリアルタイムの状況がわからないという問題点がある。ロボットや固定する場所に設置したカメラを利用すれば、現場の今の状況を取得できるが、ロボットの場合、その場で自由に移動できるのは操作者1人、カメラが固定の場合は全員が特定の一箇所しか見ることができない。さらに多くの場合、観光は1人で行くより家族や知り合いなどと複数人で行き、見たものについて感想を交わしながらより楽しめるものと考えられるが、これまでの遠隔体験システムは、ユーザ同士のコミュニケーションがほとんど考慮されていない。

2.2 遠隔コミュニケーション

一方、テレビ会議、共同作業、分散講義などの様々な分野において、複数のユーザが円滑に遠隔コミュニケーションできるシステムの研究が盛んで行われている。

Knapp[6]らはコミュニケーションを成立させる要因として以下の4項目を挙げている。

1. バーバルコミュニケーション
2. ノンバーバルコミュニケーション
3. オブジェクトへの参照
4. 環境への参照

バーバルコミュニケーションは、言語によるコミュニケーションであり、ノンバーバルコミュニケーションは、体の動き、姿勢、顔の表情や視線などの身体情報である。そしてコミュニケーションは興味の対象となるオブジェクトや、人が存在する環境に関係することもある。これがオブジェクトへの参照、環境への参照という形でコミュニケーションを成立させる。

オブジェクトや環境は、コミュニケーションのコンテキストを一致させるために重要である。例えば、環境の中にあるオブジェクトを指差して「これを見て」と言いながら、別の地点を指差して「昨日あそこに猫がいた」と相手に伝えることが、バーバル情報とノンバーバル情報でオブジェクトや環境を参照している行為である。

バーバル情報の伝達は昔から電話やメールなどを利用して実現されていたため、以下他の要因について詳細を述べる。

2.2.1 ノンバーバルコミュニケーション

ノンバーバルコミュニケーションはさまざまな研究で分類されており[7,8], ここは共同遠隔体験において特に重要だと考えている対人的空間および身体情報について役割を説明する.

1. 対人的空間

対話者たちの配置, 方向, お互いの距離が対話の生成や方向性を影響し, 会話の内容とともに変化する場合が多い.

2. 身体情報

- 体の姿勢

対象物への指示動作以外も, 相手の注意を引いたり, 感情の表現などに使われる.

- 視線

相手がどこを注視していることは, 相手が何に興味を持っていることを理解するに必要なだと考えられる.

- 表情

顔の表情は人間の感情を直接反映するものである. 電話でコミュニケーションする時に, 時々相手の感情を正しく読み取れないのは, 表情が見えないことが要因だと考えられる.

2.2.2 環境情報

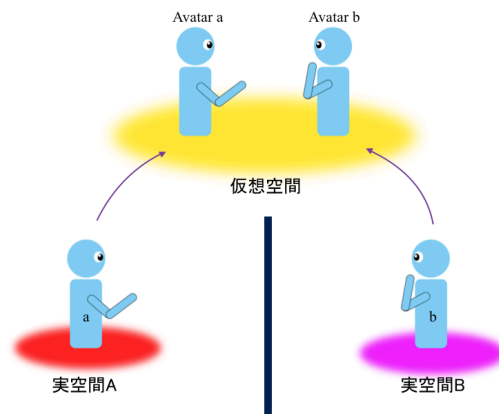
オブジェクトへの参照, 環境への参照に関して, 従来の研究では, 普通のビデオ通話では空間的不統合によって参照が難しいという問題を解決するため, コミュニケーションする複数のユーザが共同に存在する「場」を用意し, この共存空間で対話者の相対位置関係および環境との関係を再構築するアプローチを提案している.

この共存空間と双方の現場空間の関係について検討すると, 磯らが提案する Com Adapter [9]では, 図 2.5 に示すように主に 3 つのアプローチにまとめられる.

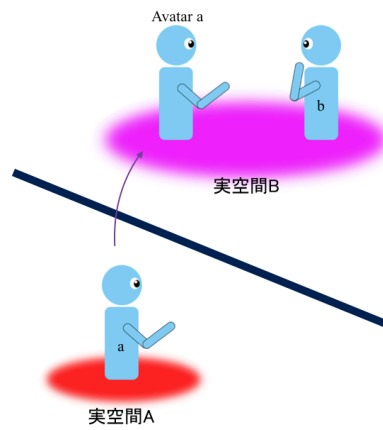
1 つ目のアプローチは, これまで多くのアプローチと同様に, 複数のユーザが実際に居る現場とは全く別の, 全てのユーザが共有するバーチャルリアリティ空間を新たに用意し, コミュニケーションするための現場とする仮想空間共有技術である. このアプローチを図 2.5(a)に示す. この方法を使用した例として, [10,11,12,13]などが挙げられる.

2 つ目のアプローチは, 一方の側だけ現場から身体を分離させ, その身体を遠隔の相手の現場に表現することである. このアプローチは, 遠隔操作ロボットなどを利用したテレプレゼンスやヘッドマウントディスプレイ内に相手の指示映像などを投影するミックスドリアリティなどの技術に見られるアプローチである. このアプローチを図 2.5(b)に示す. 例としては, [14,15,16,17]などが挙げられる.

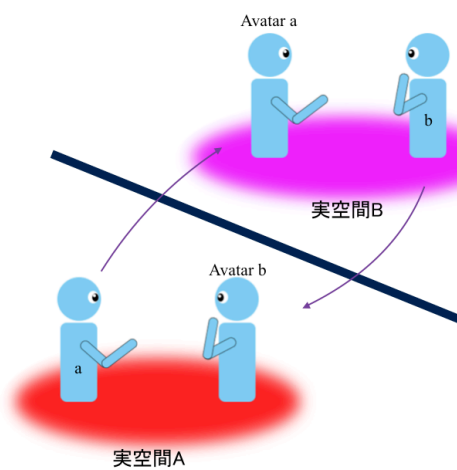
そして, 3 つ目のアプローチは, 物理的に同じ環境を用意し, 対話双方を相手の現場に身体を表現し, その中で限定したインタラクションを支援する手法である. このアプローチを図 2.5(c)に示す. [18,19]はこの方法の例である.



(a) 遠隔参加者全員が1つのバーチャル空間に表現される



(b) 遠隔参加者の片方が相手の空間に表現される



(c) 遠隔参加者がお互い相手の空間に表現される

図 2.5 遠隔コミュニケーションにおける空間の使い方

現存する遠隔体験システムの中，リアルタイム映像を提供できるシステムの視点移動ができない，またはデバイスの数が多いという問題点を解決するため，本研究は自由視点生成技術に着目した．

2.3 自由視点生成

実シーンの情報を仮想空間内に取り込み，様々な視点位置からの現実環境の見えを仮想空間内に再現する自由視点画像生成は昔から研究されている．代表的な手法として，イメージベースレンダリング (Image-Based Rendering : MBR) 法とモデルベースレンダリング (Model-Based Rendering : MBR) 法が挙げられる．以下はそれぞれについて説明をする．

2.3.1 モデルベースレンダリング法(MBR)

モデルベースレンダリング法は 3 次元形状復元手法により得られた対象シーンの 3 次元形状を利用し，仮想視点からのシーンの見えを画像として生成する技術である．この手法の特徴は，同一の 3 次元モデルを使用して生成した画像上においては，いかなる仮想視点の位置・姿勢であってもそれぞれの生成画像から復元できる幾何形状は一貫していることと，仮想環境の加工が容易であること，品質はモデルの完成度に依存することが挙げられる．MBR 法を基に生成した自由視点画像の一例を図 2.6 に示す．



図 2.6 復元した 3 次元形状を基に生成した自由視点画像例 ([20] より)

2.3.2 イメージベースレンダリング法(IBR)

イメージベースレンダリング法は、対象環境の三次元形状を明示的に復元することなく、物体やシーンを撮影した画像群を入力とし、これらを変形・合成することにより自由視点画像を生成する手法である。この手法では、対象物体の形状や反射特性を直接利用しないため、MBR法では正確な幾何形状の推定が難しい複雑な形状を持つ木などの物体に対しても、欠損の生じない自由視点画像を生成可能である。IBR法はさらに蓄積した画像群を直接変形・合成することで自由視点画像を生成する手法モーフィング[21]と、光線情報を用いる手法[22]に大別する。それぞれ生成した自由視点の画像例を図 2.7 と 2.8 に示す。しかしどの手法も、一枚の画像を生成するにたくさんの画像を入力する必要があるため、1つのカメラを用いて動的な対象を処理するのは困難である。

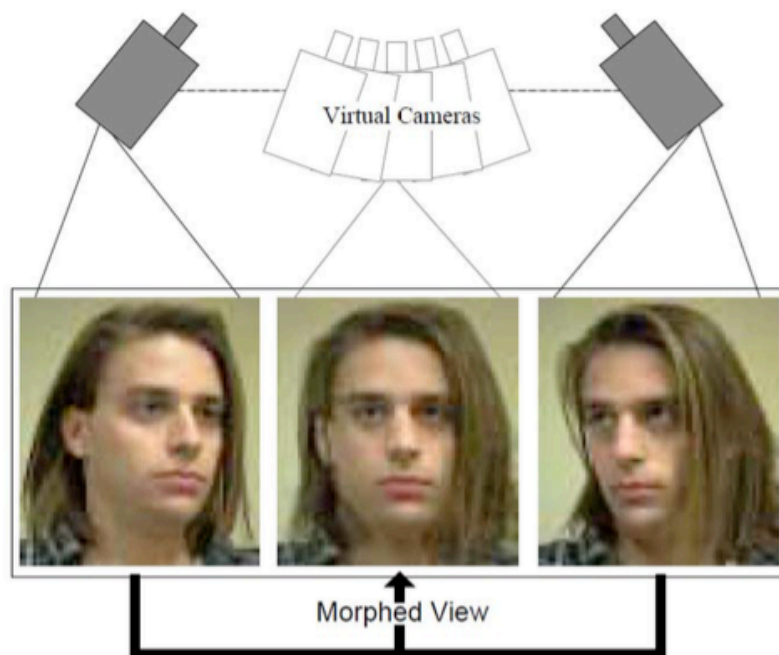


図 2.7 モーフィング([21] より)



図 2.8 ライトフィールドレンダリングで生成した自由視点画像例 ([23] より)

2.4 本研究の位置付け

本研究は現存のお遠隔体験システムの短所を踏まえて、遠隔地の映像を自由視点で見れること、ユーザ同士が円滑にコミュニケーションできることを実現するシステムの構築を目指す。

カメラを大量に使わずに遠隔地の幅広い範囲で情報を取得するため、移動ロボットを利用することを提供する。

自由視点の生成については、1つのシーンに対して、あらゆる角度の写真を取得し、その大量な写真をリアルタイムでユーザに送信するのは難しいと考えられる。本研究は移動ロボットにカメラと距離センサを搭載し、遠隔地のモデルを作成し、MBR法に基づいて自由視点を生成することを提案する。

ユーザたちに円滑なコミュニケーションを実現するため、アバタを利用することにした。コミュニケーションに必要なノンバーバル情報をアバタを用いて表現し、アバタと遠隔地の情報を1つのバーチャル空間に統合することでコミュニケーションに必要な環境情報を提供する。

第3章 遠隔共同体験システムの構成

3.1 システム概要

提案するシステムのイメージ図を図 3.1 に示す。遠隔地にロボットを設置して、周りの情報を取り込み、その情報に基づき共有するバーチャル空間に遠隔地を再築する。ユーザ用操作端末はユーザの情報を取得し、ユーザの情報に基づいて共同体験バーチャル空間にユーザのアバターを作成する。ユーザたちは自分のアバターの移動を操作しながら、アバターの視点から共同体験バーチャル空間を見ることができる。

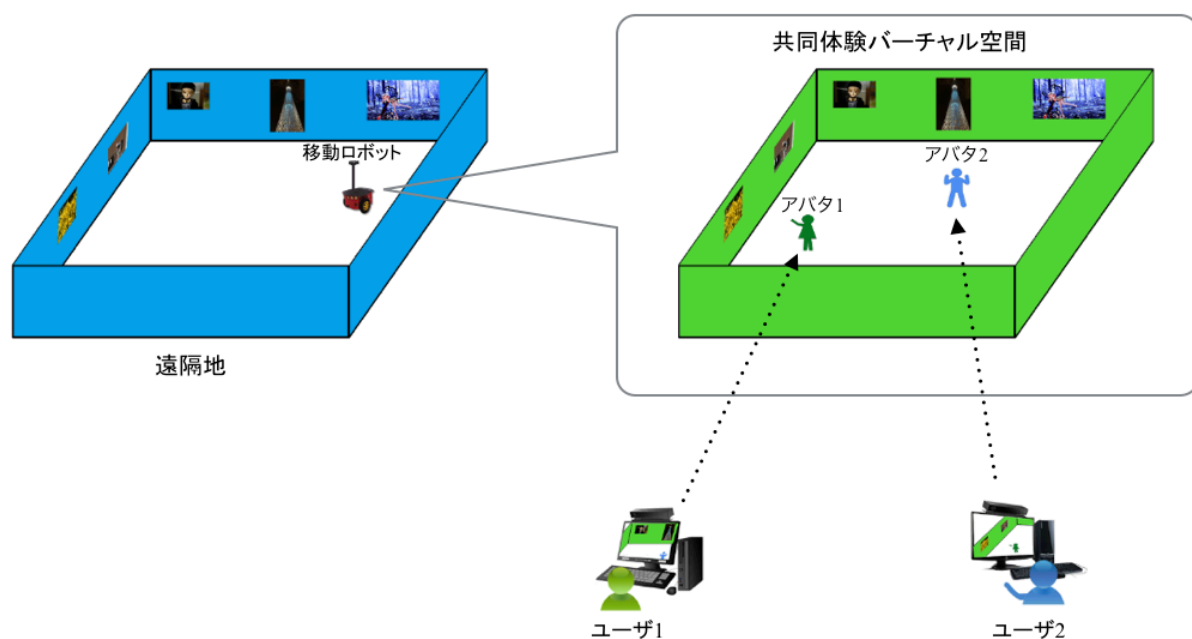


図 3.1 共同遠隔体験システムのイメージ図

3.2 システム構成

図 3.2 にシステムの構成を示す。

一度広範囲の画像を取得するため、ロボットにはカメラを3つ設置した。カメラの他、周囲障害物の距離情報を取得するためのレーザレンジファインダ、情報処理のためのPCも搭載し、映像など遠隔地のいろんな情報を各ユーザに送信する。ユーザ用操作端末側はローカルのユーザ情報を取得し送信しつつ、ロボットとリモートのユーザ情報から受け取った全ての情報を統合し、ゲームパッドを利用して視点を選択する。ユーザのインターフェースは選択した視点に合わせて、ユーザに遠隔地と同行者の情報を提示する。

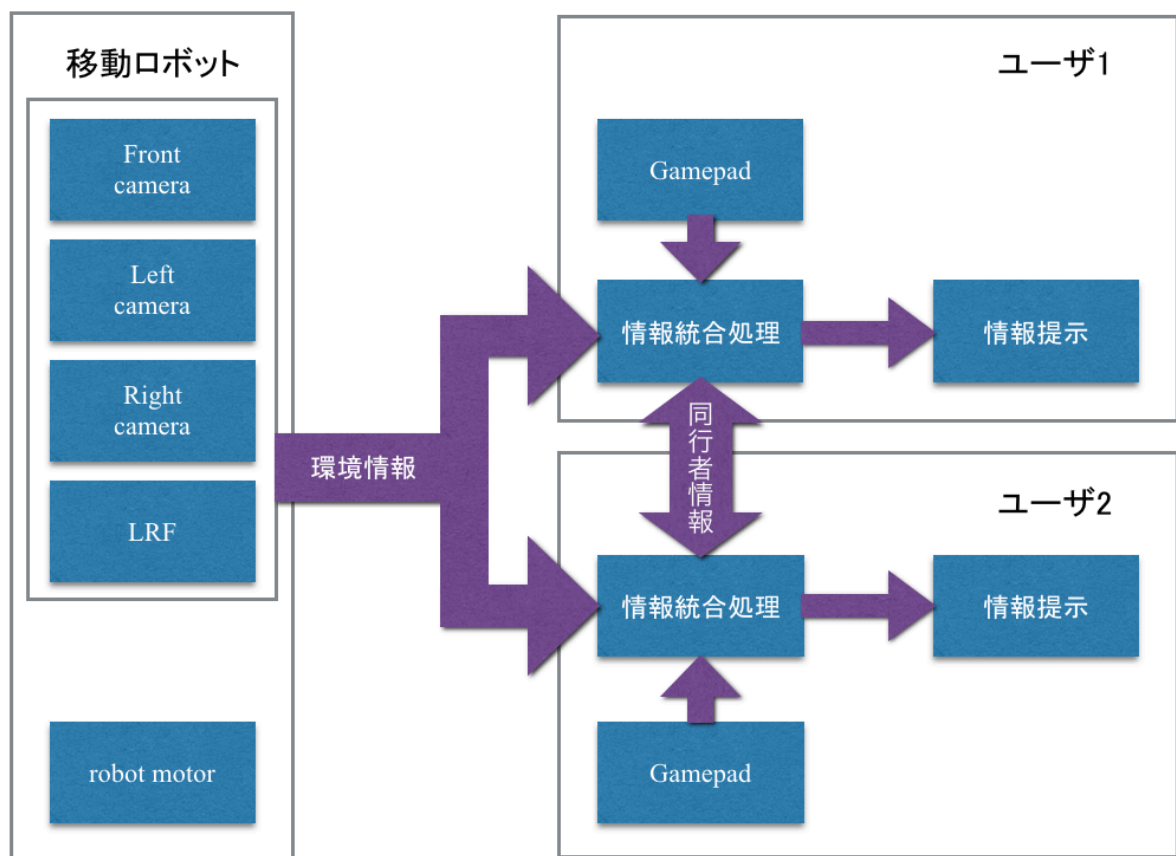


図 3.2 遠隔共同体験システムの構成

3.3 ハードウェア構成

本研究で使用する移動ロボットと、それに搭載したセンサを表 3.1 に示す。移動ロボットは2輪駆動型ロボットを使用する。ロボットには自己位置推定と周囲の歩行者をリアルタイムで認識するためのレーザレンジファインダ(以降 LRF)と、遠隔地の映像を取得するためのカメラが搭載され、センサが取得した情報を IPC(Inter Process Communication)[24]経由してユーザに送信する。ロボットの構成を図 3.3 に示す。

ユーザ用操作端末は環境情報と他のユーザ情報を統合し、共同体験バーチャル空間にすべての情報を表示させる。ユーザを表現するためのアバタにはユーザの視点位置のほか、表情を表すための顔映像、体の姿勢情報は Kinect を利用して取得する。視点の操作はゲームパッドを使用する。ユーザ側システムの構成を図 3.4 に示す。

以下、本システムで使用したそれぞれの機器について説明する。

	メーカー	型番
移動ロボット	リバスト(株)	PIONEER-3DX
カメラ	SJCAM	SJ5000
レーザレンジファインダ	北陽電機(株)	UTM-30LX
距離画像センサ	Microsoft Co.	Kinect for Windows Ver. 2.0

表 3.1 移動ロボットと搭載するセンサ

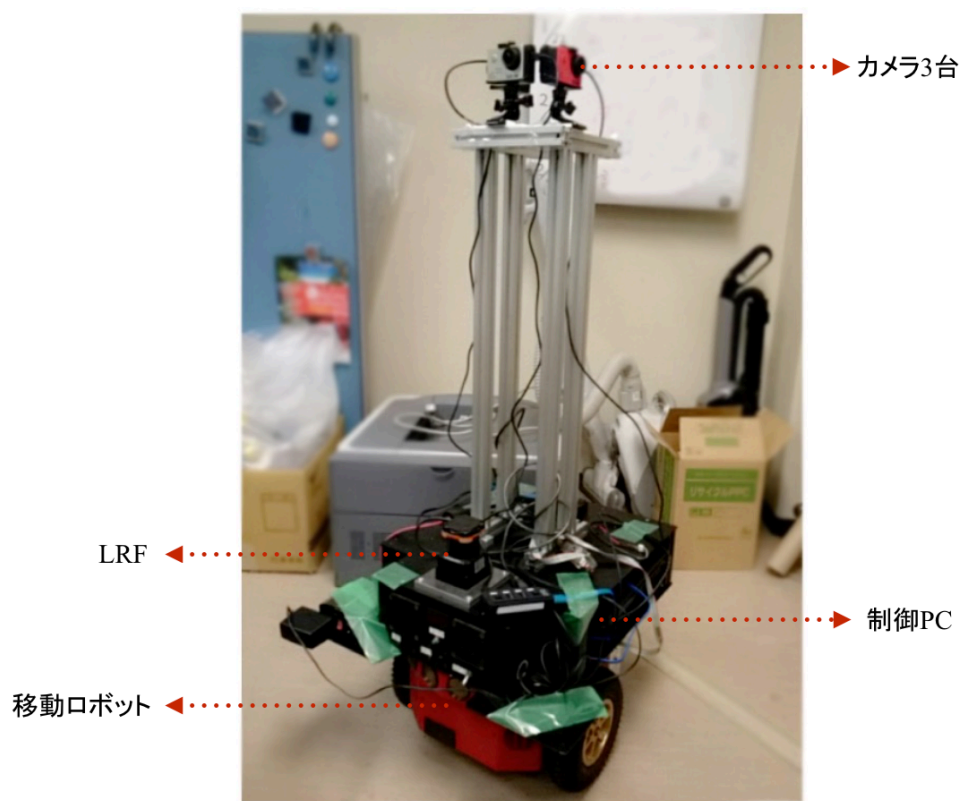


図 3.3 ロボットの構成



図 3.4 ユーザ側の構成

3.3.1 移動ロボット

移動ロボットは2輪駆動型ロボットを使用する．ユーザはロボットを意識する必要がなく，ロボットへの干渉も不可能だが，ロボットの位置に視点を縛られることもない．広い範囲で環境情報を取得するために，ロボットは一箇所に止まらず，遠隔地を適当に動き回す．

3.3.2 カメラ

カメラが取得した映像はユーザ用操作端末の共同体験バーチャル空間をリアルタイムで更新するためである．ロボットの周囲情報を幅広く取得するため，SJCAM社製の広角カメラ SJ5000 を3つ使用することにした．SJ5000の対角の画角は約170[deg]で広い範囲を1枚の画像に収められるが，画像周辺部の歪みが大きく，4.3節で述べる自由視点の生成に問題がある．そこでOpenCV 3.0のカメラキャリブレーション[25]を行い，歪みを修正する．修正した後の水平画角は約90[deg]となる．移動ロボットにカメラを3台搭載するのはLRFを1台使用した時のスキャン範囲（270[deg]）に合わせるためである．カメラの仕様を表3.2に，配置を図3.5に示す．それぞれ前，左，右に向き，画像が連続になるように位置を調整した．

	SJ5000
動画撮影	1280*720[p]
画角	170[deg]
fps	30
インターフェース	USB 2.0

表 3.2 使用されたカメラの仕様



図 3.5 カメラの配置

3.3.3 LRF

SLAM(Simultaneous Localization and Mapping)と歩行者検出するため、図 3.3 に示すようにロボットの前部に LRF を取り付けた。使用する LRF のスキャン範囲を図 3.6 に、仕様を表 3.3 に示す。LRF が検出した歩行者を画像から抽出するため、LRF のスキャン範囲とカメラの画角範囲が重なるように位置を調整した。

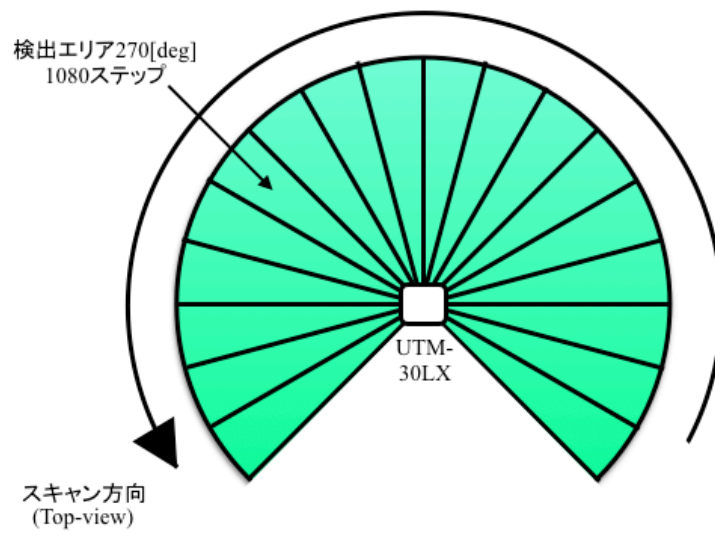


図 3.6 LRF のスキャン範囲

機械名	SJCAM-SJ5000
測距原理	Time Of Flight
測距範囲	0.1~30[m](白ケント)
角度	270[deg]
走査時間	25[ms/scan]

表 3.3 使用される LRF の仕様

3.3.4 Kinect for Windows

コミュニケーションに関して、ノンバーバル情報が重要な役割を担っていることがすでに証明され[26]、言語情報だけではなく、表情や視線、姿勢、身体動作といった様々な情報を合わせて用いることにより、より円滑なコミュニケーションを行っている。遠隔コミュニケーションの場合、空間分離によってノンバーバル情報と対話環境の不整合の問題を解決する研究はたくさん行っている[27,28,29]。特にビデオアバタ[30]を使用する場合、ユーザの立体映像をそのまま使われるため、より正確にノンバーバル情報を表現できる。しかしこの方法はアバタを立体に表現するため、いろんな方向からユーザを撮影する必要がある。

本研究はユーザが手軽に利用できるシステムを目指すため、全身像のモデルを作成せず、顔部分の映像とジェスチャ情報だけ取り込んだ CG(Computer graphics)アバタを利用することにした。表情とジェスチャの取得には Microsoft 社のデバイス Kinect を利用した。Kinect の構成を図 3.7 に、取得できる体の骨格データを図 3.8 に、スペックを表 3.4 に示す。

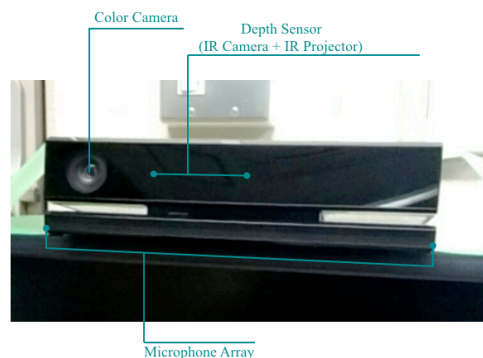


図 3.7 Kinect for Windows v2

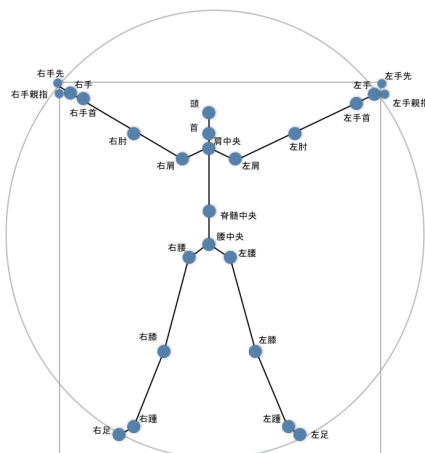


図 3.8 Kinect SDK2.0 の骨格データ

機械名	Kinect for Windows ver.2
RGB カメラ解像度	1920*1080[p]
Depth センサ解像度	512*424[p]
人物の検出範囲	0.5~4.5[m]
深度の取得範囲	0.5~8.0[m]
関節	25[関節/人]

表 3.4 Kinect for Windows v2 のスペック

3.3.5 ゲームパッド

ユーザの視点操作はゲームパッドで行う。ゲームパッドは Logitech の F310 にした。図 3.9 が示すように、ゲームパッドの操作は主に 2 つがある。右スティックは姿勢を変えずに移動，十字ボタンは姿勢の回転となる。



図 3.9 Gamepad の操作

3.4 ソフトウェア構成

システムを利用する前に環境地図を作成する必要がある。本システムは MRPT (Mobile Robot Programming Toolkit) [31]の ICP-SLAM アルゴリズムを利用し、ロボットに搭載された LRF から周りの障害物距離を使い 2 次元環境地図を作成する。

図 3.10 が示すように、システムを利用する際に、ロボットからユーザに送信するデータは 3 つのカメラ画像，SLAM により取得したロボットの現在位置，ロボットの周りにいる歩行者の位置である。また同行者間で交わす情報は顔映像および両腕ジェスチャ，同行者の視点位置と姿勢である。ユーザ用操作端末はこれらの情報を利用し，OpenCV[32]および DX ライブラリ[33]を使って共同体験バーチャル空間を立ち上げる。OpenCV はカメラ画像の処理，DX ライブラリはコンピュータグラフィックスを生成するために使われる。ロボットの現在位置推定は MRPT を利用して行う。歩行者検出はパーティクルフィルタを用いた人物追跡[34]という方法を利用する。ユーザの顔画像取得と両腕ジェスチャ取得は The Kinect for Windows SDK 2.0[35]を使用する。

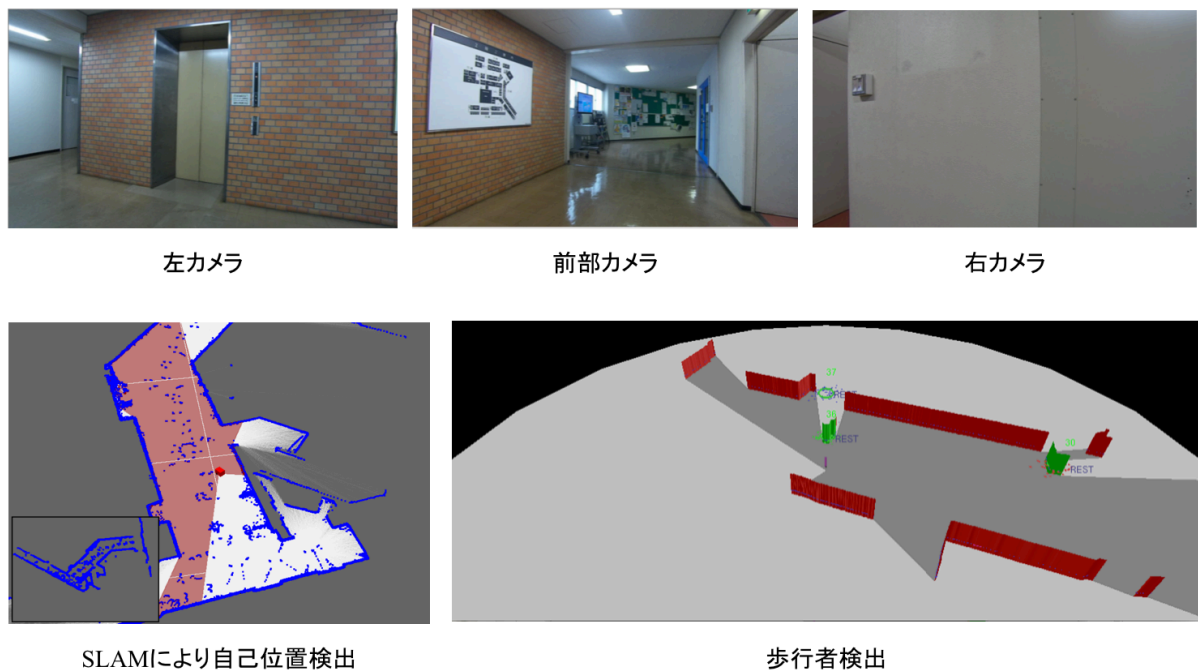


図 3.10 ロボット側がユーザに送信する情報

第4章 共同体験バーチャル空間生成のアルゴリズム

4.1 共同体験バーチャル空間の構成

共同体験バーチャル空間は以下の3つの部分によって構成される。

- 遠隔地の位置固定環境
- 遠隔地にいる歩行者
- 同行者アバタ

インターフェースの一例を図 4.1 に示す。以下それぞれの部分生成について詳細を述べる。



図 4.1 遠隔体験システムのインターフェース一例

4.2 共同体験バーチャル空間生成の基本的な流れ

実環境情報と同行者情報を共同体験バーチャル空間に取り込むための処理の流れを図 4.2 に示す。以後、位置固定環境、歩行者、同行者の情報処理をそれぞれ詳細について述べていく。

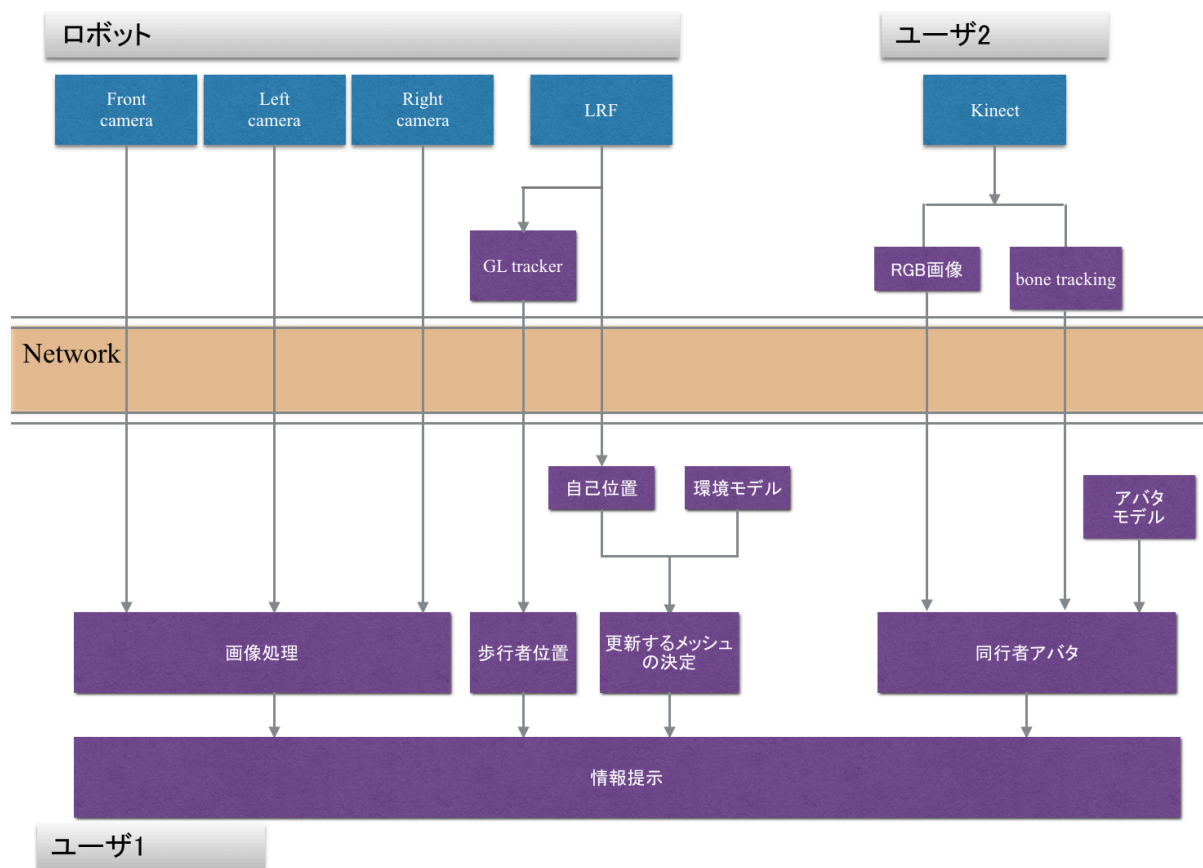


図 4.2 インターフェース生成の流れ

4.3 位置固定環境の生成

本研究は一箇所で取得したカメラ画像から、あらゆる視点位置の画像を生成するために、MBR 法に基づき、事前取得した 2 次元地図に地面と天井の高さを加え、位置固定環境モデルを作成した。環境が複雑な場合はさらに正確なモデルが必要だが、本研究の研究範囲には含まれないため省略とする。

4.3.1 位置固定環境生成の流れ

位置固定環境生成の流れを図 4.3 に示す。以後それぞれの部分について詳細を述べていく。

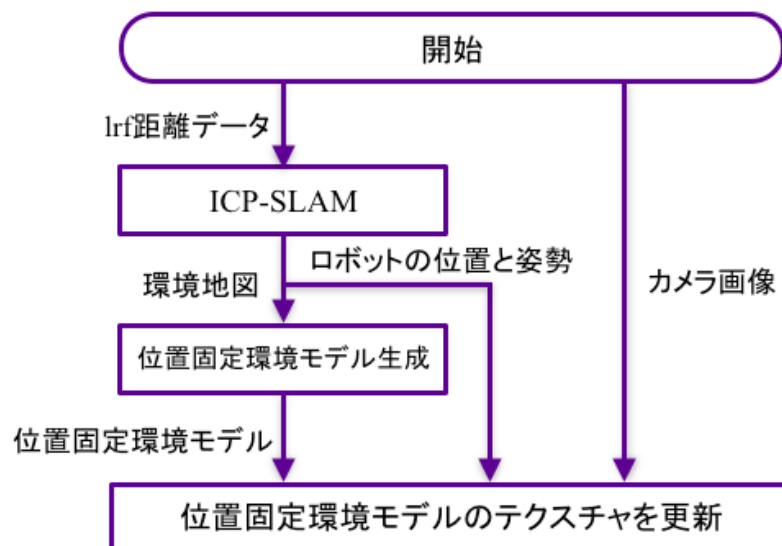
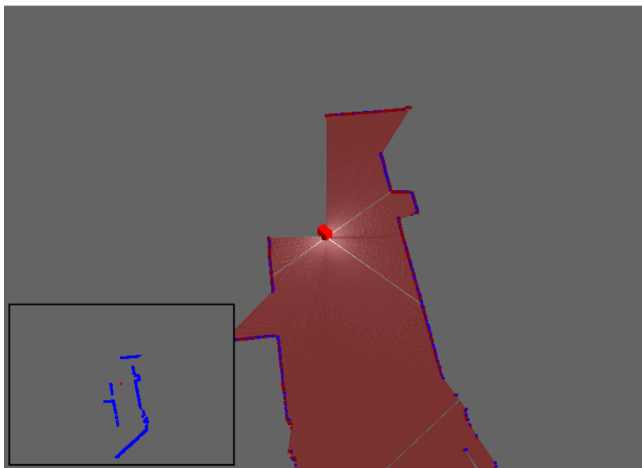


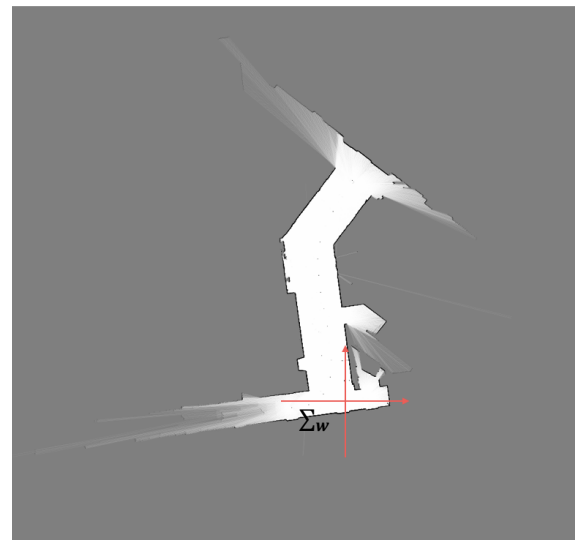
図 4.3 位置固定環境生成の流れ

4.3.2 SLAM を用いた環境地図と自己位置推定

本システムで取扱う環境モデルは壁の部分のみとする．環境モデルを生成するには環境地図を作成した．地図作成には ICP-SLAM(Iterative Closest Point - Simultaneous Localization and Mapping)[36]を利用した．ICP-SLAM とは LRF から得られる周囲の距離情報（図 4.4(a)）を連続の 2 ステップ分のデータを使用して，それらの点群が一番マッチングする位置を計算し，ロボットの現在位置と姿勢を推定しながら，同時に地図を生成するアルゴリズムである．作成した地図を図 4.4(b)に示す．初回走行時の移動ロボットの初期位置姿勢をワールド座標系 Σ_w の原点とし，SLAM により移動ロボットの現在の位置 ${}^w p_r$ と姿勢 ${}^w R_r$ が推定される．



(a) LRF が取得した周囲の距離情報



(b) 地図生成結果

図 4.4 ICP-SLAM により環境地図を作成

実空間を再築するには位置固定環境の3次元位置を知る必要がある．その位置に正確なモデルを作ることであらゆる視点から正しい形状を見ることができる．本研究は天井から地面までの壁を位置固定環境とし，環境モデルを生成する．環境モデルはポリゴンメッシュにより構成される．ワールド座標系 Σ_w は高さ方向を Y 軸とし， Σ_w およびカメラ座標系 Σ_c におけるカメラの高さを $Y=0$ とし，カメラの高さ h は固定で既知，天井から地面までの距離 H も既知のため，地面の Y 座標は $-h$ ，天井の Y 座標は $H-h$ である．環境地図の平面に地図を短い線分の集合に近似し，1つ線分に地面と天井の高さを加え，1つの3次元ポリゴンメッシュを生成する．環境地図の生成はICP-SLAMにより行い，作成した地図データには，LRFの計測誤差が原因となるノイズがたくさん残っているため，モデルを生成する前にノイズと思われる部分を削除し地図を修正する．生成した位置固定環境モデルを図4.5に示す．縦の黒い線はポリゴンの辺を表す．

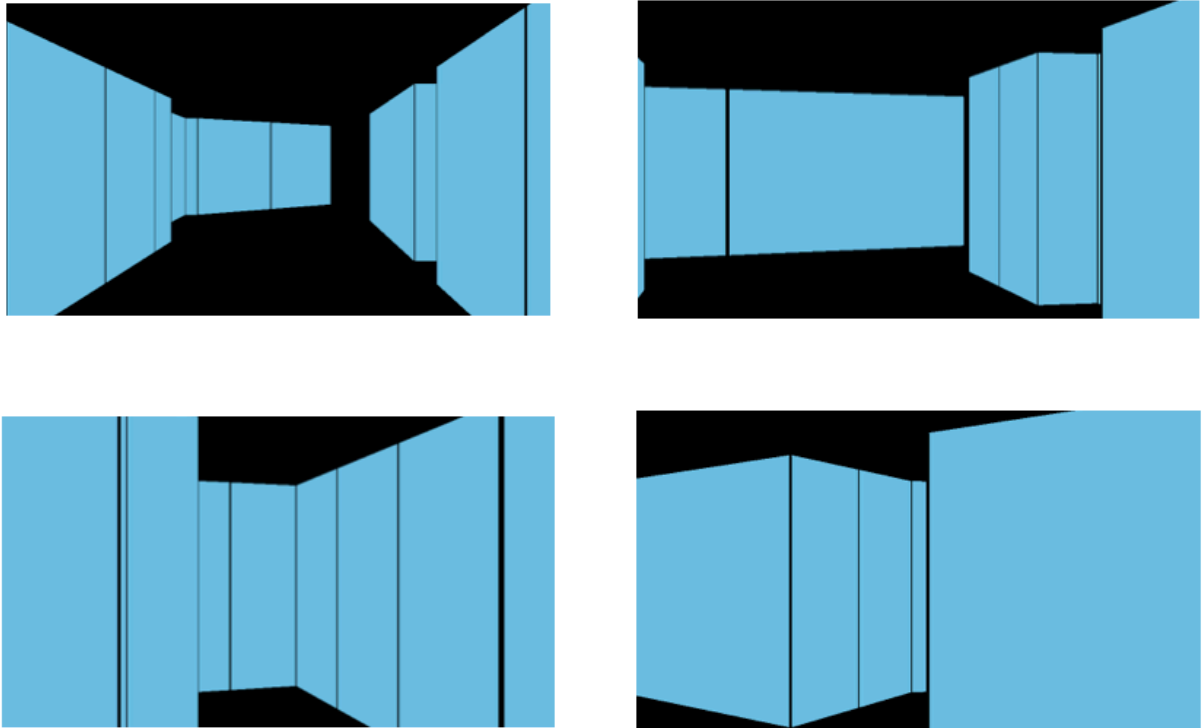


図 4.5 位置固定環境のモデルの生成結果

4.3.4 テクスチャマッピング

カメラ画像を位置固定環境モデルにマッピングするには, カメラキャリブレーションを行い, カメラ画像の歪みを取り除き, 環境モデルのワールド座標系 Σ_w の3次元座点とカメラ画像平面との対応関係を計算する必要がある. カメラのピンホールモデルを図4.6に示す.

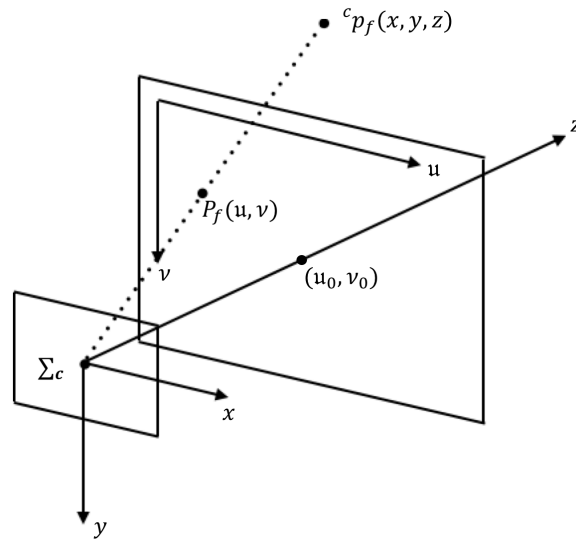


図 4.6 カメラのピンホールモデル

画像面への射影変換には以下の式が成立する

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

略して

$$sm = A[R/t]M \quad (2)$$

ここで ${}^c p_f = (x, y, z)$ はワールド座標系 Σ_w の3次元座標を表し, $P_f = (u, v)$ は画像平面に投影さ

れた点の座標を表す. また $m = (u, v, 1)$ とする. $A = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ はカメラの内部パラメータと

呼ばれ、カメラキャリブレーションを行うことにより算出することができる． f_x と f_y はそれぞれ X 軸と Y 軸の焦点距離， (c_x, c_y) は Σ_c の Z 軸と画像面の交点である．並進-回転の同次変換行列である $[R/t]$ は外部パラメータ行列，もしくはカメラのビュー行列と呼ばれ， Σ_w における点 $M = (X, Y, Z)$ を Σ_c に変換するための行列である．

歪みがない場合， Σ_c の点 ${}^c p_f = (x, y, z)$ が画像平面に投影した点 $P_f = (u, v)$ は

$$x' = x/z \quad (3)$$

$$y' = y/z \quad (4)$$

$$u = f_x * x' + c_x \quad (5)$$

$$v = f_y * y' + c_y \quad (6)$$

となる．

ところが実際のカメラレンズは歪みを持つてるので，上式を下のように拡張する．

$$x' = x/z \quad (7)$$

$$y' = y/z \quad (8)$$

$$x'' = x'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2d_1 x' y' + d_2 (r^2 + 2x'^2) \quad (9)$$

$$y'' = y'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2d_2 x' y' + d_1 (r^2 + 2y'^2) \quad (10)$$

$$r^2 = x'^2 + y'^2 \quad (11)$$

$$u = f_x * x'' + c_x \quad (12)$$

$$v = f_y * y'' + c_y \quad (13)$$

ここで k_1 ， k_2 ， k_3 は半径方向の歪み係数， d_1 ， d_2 は円周方向の歪み係数である．カメラキャリブレーションを行い，これらの歪み係数を求めることができる．つまり点 P_f の歪みがない場合の座標と，歪みが発生した場合の座標の相対関係が求められる．この結果を利用して画像の歪みを取り除くことができる．カメラキャリブレーションはOpenCVを利用して行い，歪み修正前と後を図4.7に示す．



図 4.7 カメラ画像の歪み修正

続いて外部パラメータ行列 $[R/t]$ を求める．ICP-SLAM を行うことにより LRF の位置 ${}^w p_l$ と姿勢 ${}^w R_l$ を取得する．LRF とカメラとの相対位置は固定するため，LRF 座標系 Σ_l におけるカメラの位置 ${}^l p_c$ と姿勢 ${}^l R_c$ は既知である．次のビュー変換式で Σ_w におけるカメラの位置 ${}^w p_c$ と姿勢 ${}^w R_c$ が求められる．

$${}^w p_c = {}^w p_l + {}^w R_l \quad {}^l p_c \quad (14)$$

$${}^w R_c = {}^w R_l \quad {}^l R_c \quad (15)$$

ここで回転行列 ${}^l R_c$ は

$${}^l R_c = \begin{bmatrix} \cos \alpha_c & -\sin \alpha_c \\ \sin \alpha_c & \cos \alpha_c \end{bmatrix} \quad (16)$$

と表せ， α_c はカメラと LRF の相対姿勢角度である．次に Σ_w におけるとある点 Z の位置 ${}^w p_z$ から Σ_c における位置 ${}^c p_z$ を計算できる．

$${}^w p_z = {}^w p_c + {}^w R_c \quad {}^c p_z \quad (17)$$

つまり

$${}^c p_z = {}^w R_c^T ({}^w p_z - {}^w p_c) \quad (18)$$

式(18)と式(5,6)から点 A の画像平面での座標 (u, v) が求められる。

この手順に従って位置固定環境モデルのポリゴンメッシュの四つの頂点が各カメラ座標系における座標および各画像平面での座標が計算できる。しかし実際は地図上の全てのポリゴンがカメラ画像に映せるではなく、別のポリゴンにより遮られたポリゴンと、距離が遠くて画像が小さすぎるポリゴンを更新する必要がない。したがってマッピングを行う前に、どのポリゴンを更新するかをまず決める必要がある。

更新するポリゴンはまずロボットから離れすぎないという条件を設け、この距離以内にいるポリゴンだけに隠面判定を行う。隠面判定は2次元環境地図を利用して行う。判断したいポリゴンの元となる線分の頂点とカメラが結ぶ直線が、2次元環境地図との交点が線分上にあると、ポリゴンが遮られてると判断し、更新するデータから排除する。また、同じ計算方法で歩行者のポリゴンとの関係も計算する。歩行者情報の取得は次節で述べる。残ったポリゴンだけ更新を行う。

続いてどのポリゴンにどのカメラ画像を使用するかを決める。まずポリゴンの元となる線分の頂点が3つのカメラ座標系における座標をそれぞれ求め、図4.8が示すように、カメラ画像面と垂直する直線との夾角が画角の半分以上であれば、このカメラの画像を使うことになる。また1つのポリゴンでも、線分の両端が別々のカメラ画像に属する可能性がある。この場合、左の画像面と垂直する直線との夾角が画角の半分のところに、ポリゴンを分断し、それぞれ別のカメラ画像から対応するテクスチャを抽出し合成する。

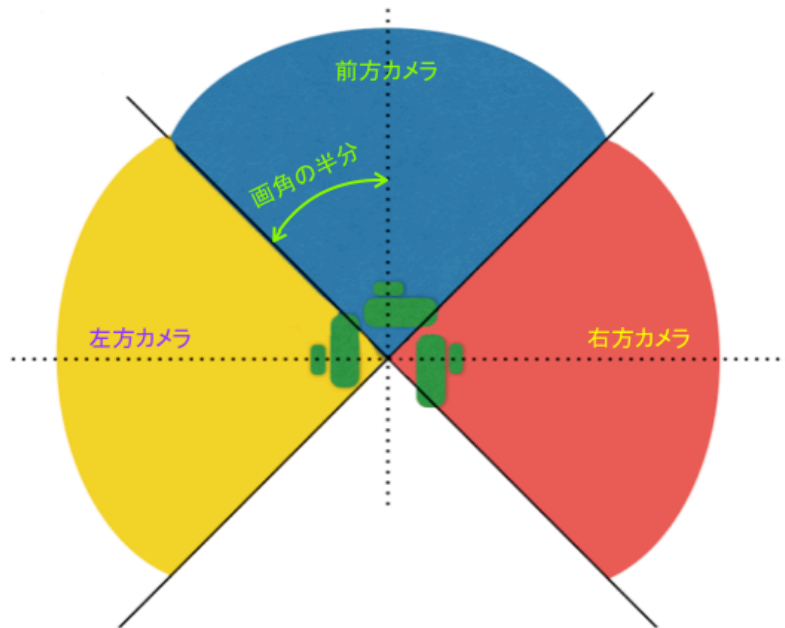


図 4.8 カメラ画像の選択

更新するポリゴンと対応するカメラ画像が決まったら、ポリゴンの4つの頂点が画像面での座標を求める。ポリゴン自体は長方形であるが、画像面に投影すると不規則な四辺形になってしまう。DX ライブラリ自体は透視変換する機能がないため、マッピングする前に OpenCV を用いて透視変換を行い、テクスチャを本来の形状に戻す。

以上の流れでカメラ画像を修正し、位置固定環境の各ポリゴンの頂点の画像面座標を求め、画像からポリゴンに対応する四辺形をポリゴンの形状に変換し、ポリゴンのテクスチャとして貼る。生成した位置固定環境の自由視点画像の例を図 4.9 に示す。システムの初期状態は前回構築した最後の記録を利用して地図全範囲のモデルを生成し、ロボットが更新せずともユーザーに全範囲の情報を得ることができる。

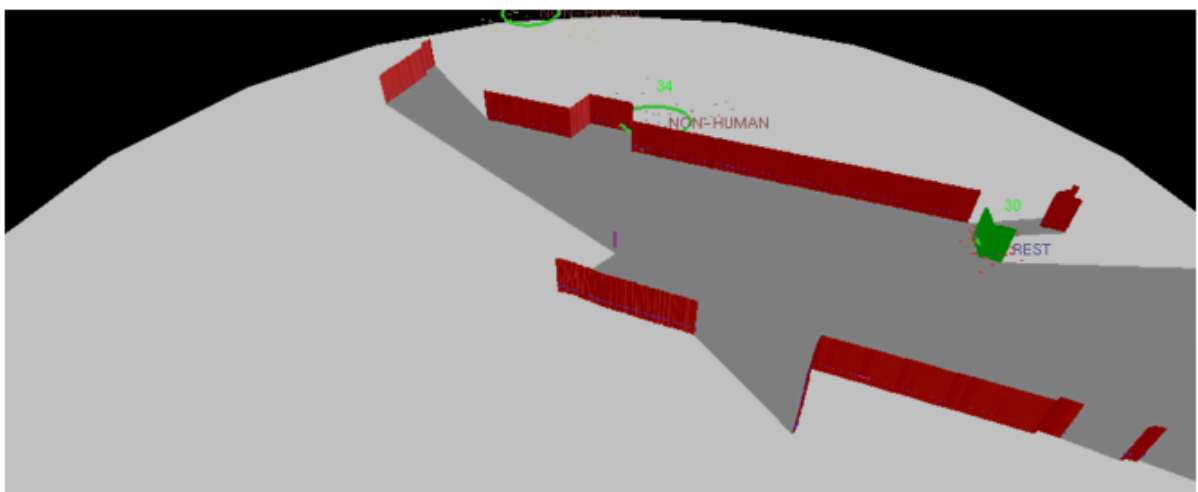


図 4.9 位置固定環境の自由視点生成結果

4.4 歩行者の描画

ロボットの周囲の人を検出するため、パーティクルフィルタを用いた人物追跡[34]をおこなう。図 4.10 にその一例を示す。これは、LRF から得られる距離情報を基に人の足と思われる対象を推測し、その重心を人の座標として得る。またこの方法は周りの障害物を人間だと誤認識する可能性もあるため、検出された歩行者座標とあらかじめ作成した地図データと比べ、壁に近い歩行者を排除し、残ったデータだけを使うことにする。

歩行者のモデルは、検出された座標を中心に、画像面と並行する平面を使う。どのカメラ画像を使うのは、前節で説明した位置固定環境ポリゴンメッシュに使う方法と同じである。また歩行者を身長 180 センチ、幅 70 センチと仮定し、この大きさのモデルを作成した。モデルのテクスチャマッピングは位置固定環境ポリゴンのマッピングと同じ方法で行う。図 4.11 に歩行者の生成結果を示す。



壁に近い検出結果を排除し、
描画する歩行者を抽出する

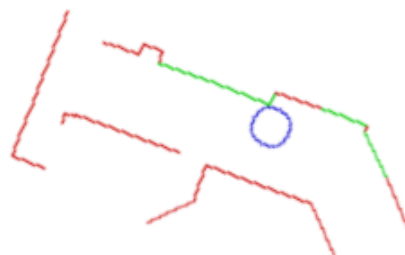


図 4.10 LRF による人物追跡

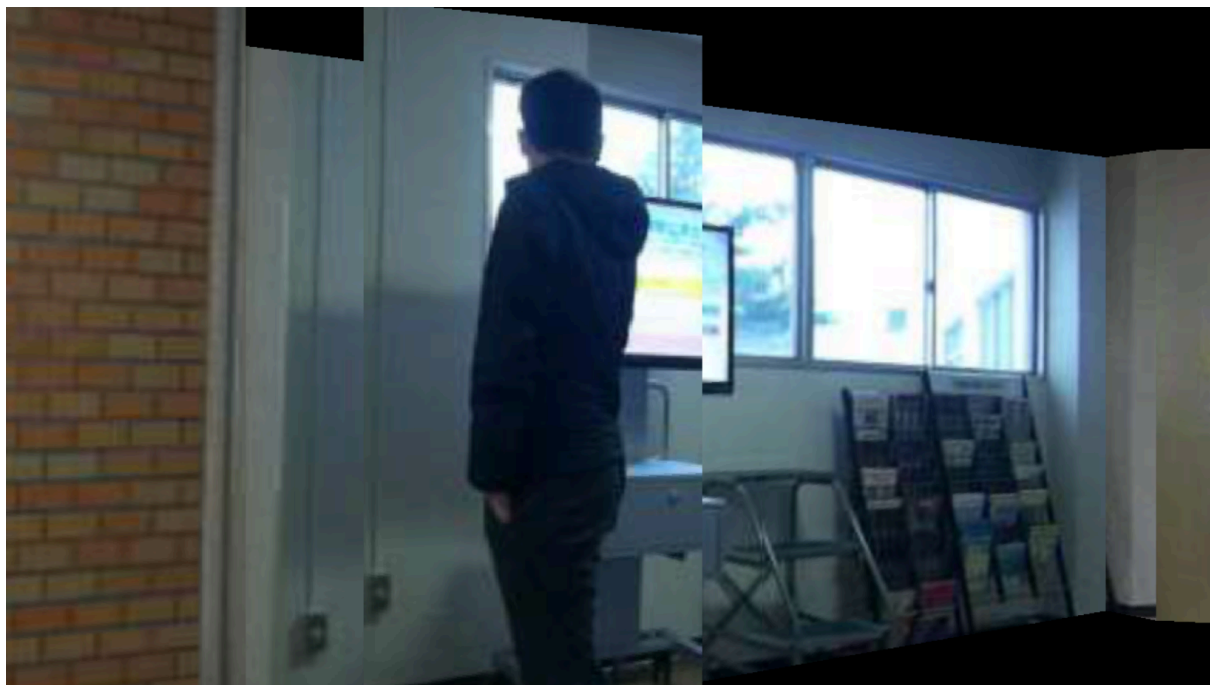


図 4.11 歩行者モデルの生成結果

4.5 同行者の描画

同行者のアバタは MikuMikuDance 用モデル[37,38]をベースに，同行者の視点位置に描画される．またアバタの向きも同行者の視点向きと一致させ，同行者が何を注目しているかをアバタの位置と向きにより表現できる．

アバタの構成を図 4.12 に示す．顔部分を透明化し，同行者の顔映像を貼り付ける．また前後を区別するため後ろ髪を残した．両腕は同行者の両腕ジェスチャを代入した．また場所移動する際には，自然に見せるため足に歩くアニメーションを代入した．

同行者顔映像と姿勢の取得に Kinect2.0 の SDK を利用した．Kinect のモーションキャプチャで得られるデータは Kinect の Depth 座標系における人体の 25 個の key Joint の座標で構成されたもの．対話する時に使うジェスチャは主に腕の部分と考え，本システムは両腕のデータだけアバタに反映し，頭の Joint の座標から Kinect の Color 座標系の座標を求め，この座標を中心に一定の範囲を抽出し，同行者の顔映像としてアバタに張り付く．

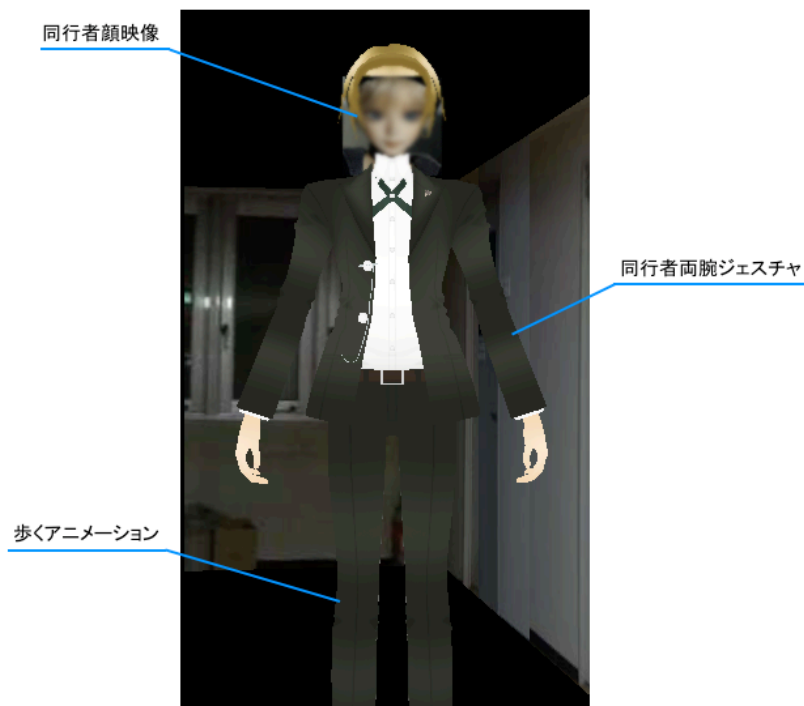


図 4.12 同行者アバタの構成

第5章 実験

5.1 実験環境

本システムのユーザビリティを評価するため、比較実験を行った。実験は学校の廊下で行った。廊下の平面図を図 5.1 に示す。

実験における動作環境を表 5.1 に示す。ロボット側にはカメラ, LRF から情報を所得し, SLAM を行うための Windows PC を搭載されている。歩行者検出のプログラムは Unix 系の OS で動くため, LRF のデータは SLAM に使われると同時に, IPC を通じて OS X PC に伝送されている。使用するライブラリのバージョンを表 5.2 に示す。

ユーザ側の実験風景を図 5.2 に示す。

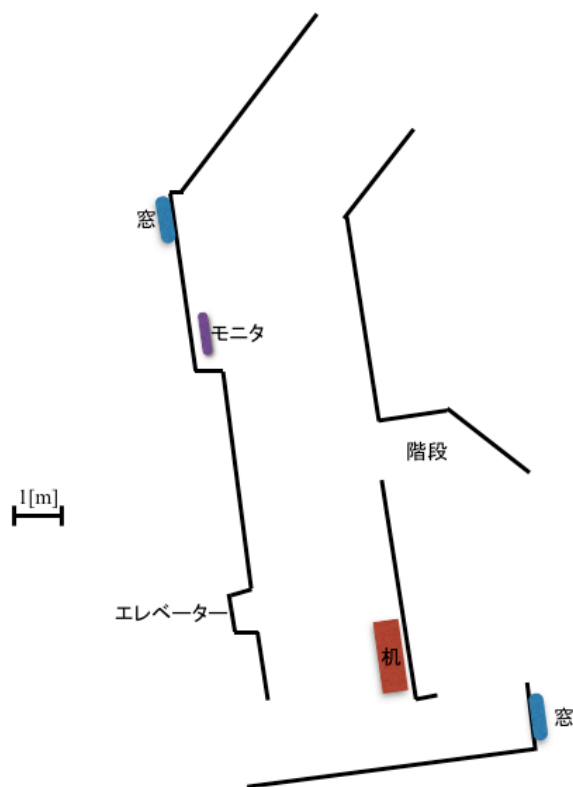


図 5.1 実験環境の俯瞰図

機械名	種類	仕様	その他
ロボット側 PC	OS	Windows 8 64bits	
	CPU	Intel Core i7-4700MQ	
	RAM	8GB	
サーバ PC	OS	OS X EI Capitan	IPC サーバ 歩行者検出
	CPU	Intel Core i5 2.40GHz	
	RAM	8GB	
ユーザ側 PC	OS	Windows 8 64bits	ユーザ 1
	CPU	Intel Core i7-3770K 3.50GHz	
	RAM	8GB	
	GPU	NVIDIA GeForce GTX650M	
ユーザ側 PC	OS	Windows 8 64bits	ユーザ 2
	CPU	Intel Core i7-2630QM 2.20GHz	
	RAM	4GB	
	GPU	NVIDIA GeForce GT 540M	

表 5.1 動作環境

SDK	用途	バージョン
MRPT	SLAM	1.3.2
Aria	ロボットの移動制御	2.9.0-1
IPC	データ受送信	3.8.6
OpenCV	画像処理	3.0.0
DXlib	CG 生成	3.13d
The Kinect for Windows SDK	ユーザ情報取得	2.0_1409

表 5.2 使用したライブラリ



図 5.2 ユーザ操作風景

5.2 実験の条件

比較実験には、アバタの効果検証，環境情報更新の効果検証，及び自由視点の効果検証するため，4つのシステムを用意した．各システムの違いを表 5.3 に示す．以下それぞれのシステムについて概要を述べていく．

実験条件	視点方式	自由視点の有無	環境情報更新の有無	アバタの有無
システム 1	第 1 人称	あり	なし	あり
システム 2	第 1 人称	あり	なし	なし
システム 3	第 3 人称	なし	あり	あり
システム 4 (提案手法)	第 1 人称	あり	あり	あり

表 5.3 実験条件の設定

5.2.1 システム 1: 共同体験バーチャル空間を更新しない遠隔体験

システム 1 では提案するシステムの位置固定環境をあらかじめモデルを生成し、一度作ったモデルは一切更新を行わず、歩行者情報も取得せず、つまり遠隔地にいるロボットは必要としないシステムである。同行者については提案システムと同じ仕組みとなっている。このシステムではユーザはそれぞれ自由に動けるため、同行者が視野範囲にいないことは多い。そこで常に同行者の位置を把握させるためにミニマップをユーザに提示した。ミニマップには環境地図、自分の視点位置、同行者の位置が描かれている。ミニマップの一例を図 5.3 に示す。



図 5.3 自由視点あり場合のミニマップ

5.2.2 システム 2: アバタを利用しない遠隔体験

このシステムはシステム 1 とほぼ同じものだが、システム 1 との唯一の違いは、共同体験バーチャル空間に、利用者のアバタを生成しないことである。

5.2.3 システム 3: 自由視点を生成しない遠隔体験

システム 3 は自由視点を生成せず、遠隔地の映像を背景として共同体験バーチャル空間に貼り付ける。このシステムは視点変更ができないため、ユーザは第 3 人称視点でアバタを操作することとなる。視点の位置姿勢はカメラのワールド座標 Σ_w における位置と姿勢である。ユーザはアバタの移動操作だけを行う。ゲームパッドのスティックの倒し角度にアバタは移動する。この角度はカメラ座標系 Σ_c における角度である。視点の操作、つまりロボットの移動操作は実験者が行う。このシステムで生じやすい問題は、複数のユーザがいるため、それぞれ違う方向に行きたい場合がある。視点は 1 つしかないため、移動ロボットの操作者はユーザの様子を見て、そのうち 1 人のアバタを追尾するしかない。システム 3 のインターフェースの一例を図 5.4 に示す。



図 5.4 システム 3 のインターフェース例

5.2.4 システム 4: 実環境の情報を基づき共同体験バーチャル空間を更新する遠隔体験

システム 4 は本研究の提案システムである。初期状態はシステム 1 のようになっている。ただしロボット PC から遠隔地の映像および歩行者情報を受け取り、それに基づいて共同体験バーチャル空間の情報を更新していく。またシステム 1 と同じく同行者の位置を把握させるためミニマップが使うことにした。ロボットの操作は実験者が適当に行った。

5.3 実験の手順

被験者は 20 代から 30 代の学生 6 名で行った。被験者に最初実験に関するマニュアルを渡し、さらに視点の操作、タスクについて口説明を行った。本システムはユーザが最低 2 人が必要であり、被験者を 2 名ずつにし、3 つのグループに分けた。また本システムの利用に関して、具体的なタスクを設けた。タスクの内容は出来るだけ早く遠隔地に貼ってあった色紙を探し出すことである。コミュニケーションのしやすさを評価するために、被験者はお互い情報交換をしてもらい、先に見つけた人は同行者に場所を教えていただける。色紙は見つかりやすくため 3 つの色違い色紙を並べて使う。

このタスクは、アバタ有無の違い、自由視点有無の違い、および第 1 人称視点と第 3 人称の違いによって、遠隔体験にもたらす影響を検証するため、ユーザたちに積極的移動とコミュニケーションをするようにと誘導した。

1 つグループの被験者 2 人はお互い見えさせないため違う教室に座ってもらい、また遠隔地も被験者が見えない場所にした。

5.3.1 実験 1

実験は 2 回に分けて行われた。以下 1 回目の実験手順を説明する。

システム 1 ではタスクを設けず、被験者たちに特に指示も出さず、10 分間自由に使用してもらった。

システム 3 は最初に被験者がシステムを慣れるまで練習操作を行い、そのあとにタスクの内容を伝えた。タスクが始まる前に色紙をロボットの初期視野外の遠隔地のどこかに貼り、実験開始から 2 人とも色紙を見つけるまでの時間を記録した。

システム 4 はシステム 3 と同じ、初期視野に色紙がない状態から始め、を見つけるまでの時間を記録した。まだシステム全体はシステム 1 とほぼ変わらないため、練習時間を省けた。色紙はシステム 3 と違う場所に配置するようにした。

実験の順番はランダムだが、システム 4 がシステム 1 より前の場合、タスクが始まる前にシステム 4 でも練習操作のフレーズを入れた。グループごとにシステム 3、4 でのロボット、色紙の初期配置をそれぞれ図 5.5(a)～(c)に示す。



(a)



(b)

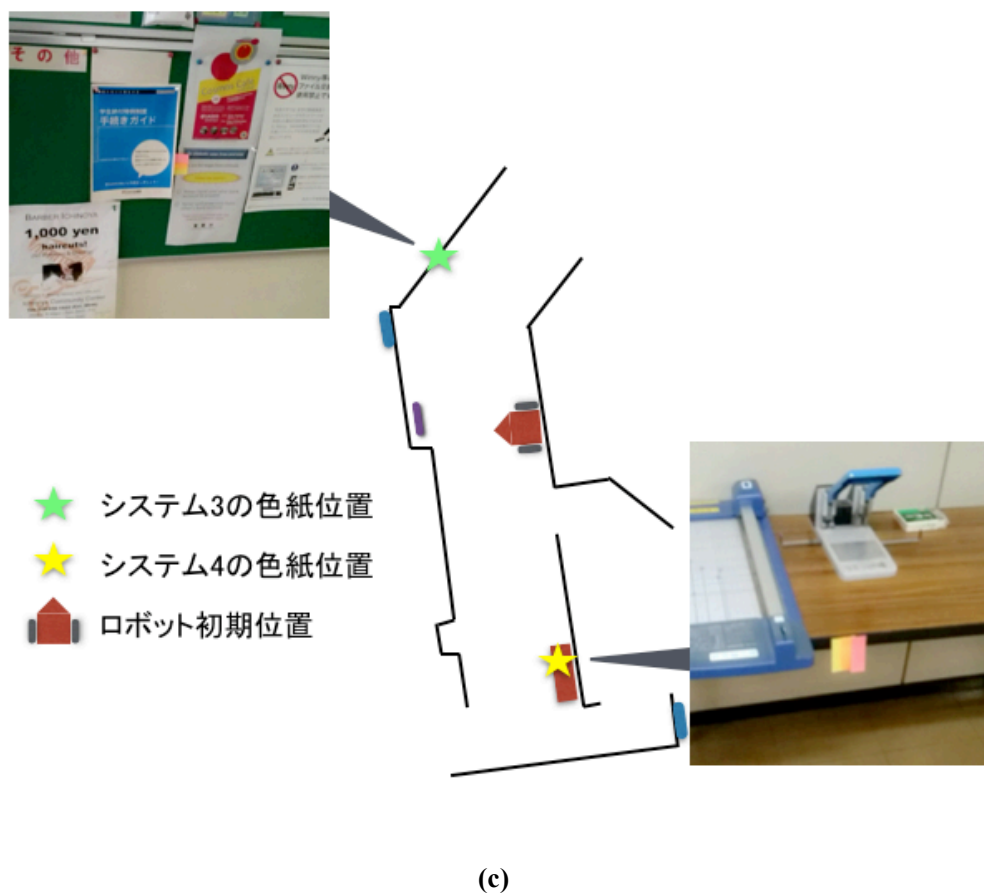
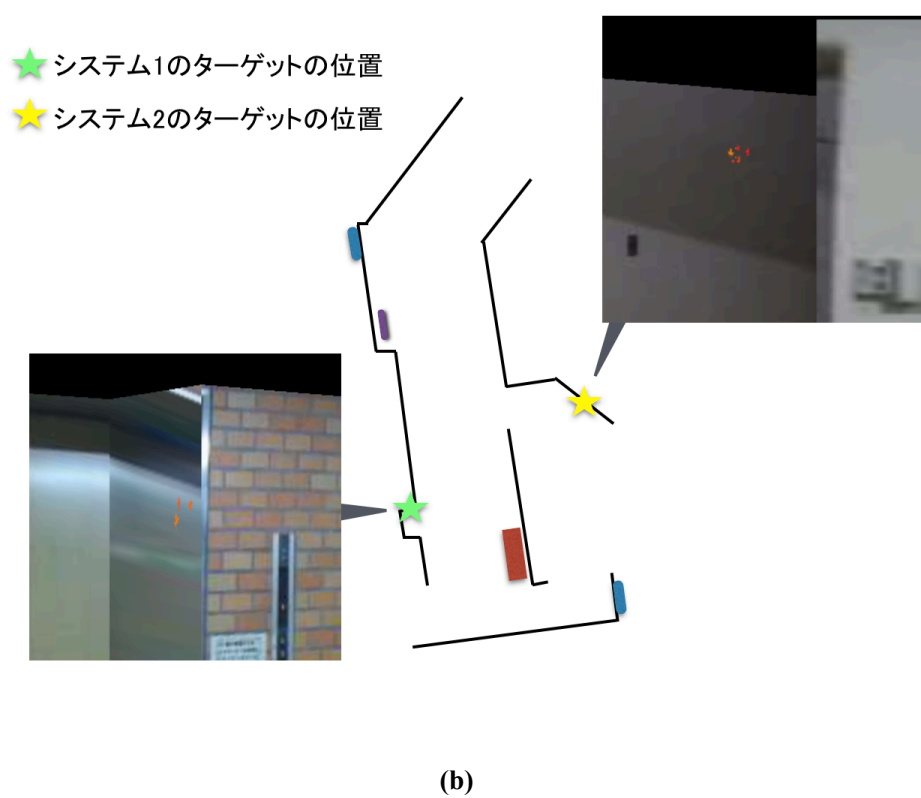
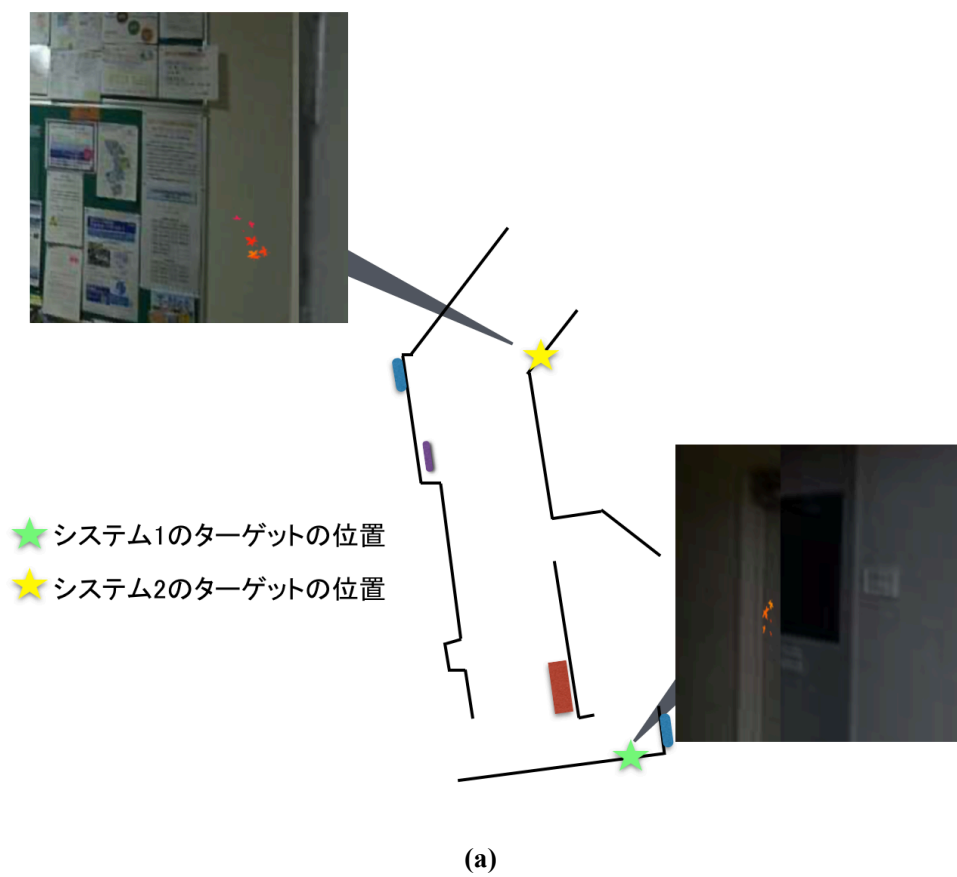


図 5.5 実験タスクの色紙の初期配置

5.3.2 実験 2（追加実験）

アバタの使用により，遠隔共同体験への影響を明らかにするため，追加実験を行った．追加実験はアバタの有無以外，全ての条件が同じようにしたシステム 1 とシステム 2 の間で行われた．

まずシステム 1 を用いて操作練習を行い，その後前回の実験と同じもの探しというタスクを設けた．タスクを達成できたら，ターゲットの位置を変更してシステム 2 の仕様でもう一度タスクを行われた．各グループのターゲット位置配置を図 5.6 に示す．



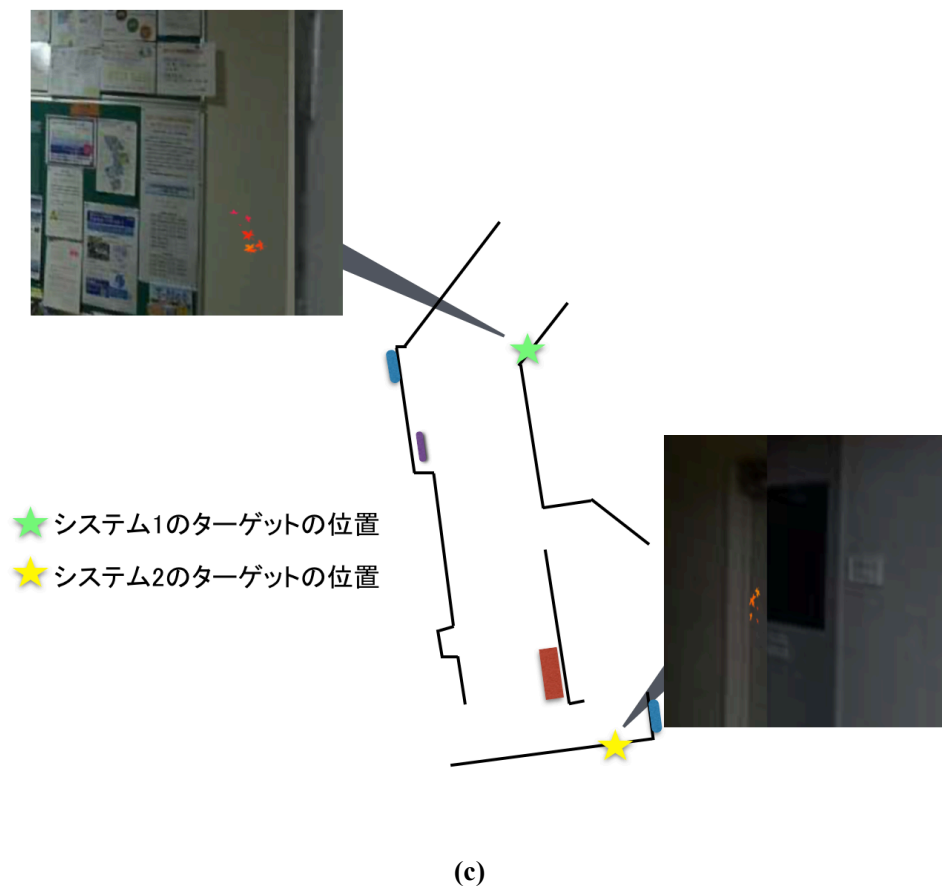


図 5.6 追加実験のターゲットの初期配置

5.4 実験の評価方法

システムの評価は実験でのタスク完成時間及び実験後のアンケートによって行った。タスクの設定は各被験者にシステムの違いを理解させるために設けたものである。

実験後、各被験者にアンケートを答えてもらい、リッカート尺度を使って各項目を評価した。質問項目の設定は、細谷の論文[39]で遠隔コミュニケーションのインターフェースを評価する項目を参考に設けた。

項目はそれぞれ臨場感、自由度、操作感、視線理解容易性、会話容易性、同室感、総合的利用意向の7つの観点に絞り、合計14問に決定した。各質問を下に挙げる。

臨場感

1. 遠隔地の今の様子が分かりましたか。
2. 遠隔地を見ている感じがしましたか。

自由度

1. 現場にいるように自由に見渡せましたか。
2. 自分の行動に制限が多いと感じがしましたか。

操作感

1. 目的地への移動は容易にできましたか。
2. 遠隔地を見渡すための操作は容易にできましたか。

視線理解容易性

1. 相手がどこを見ているか分かりましたか。
2. 相手と目が合っていると感じましたか。

対話容易性

1. 相手とやり取りしやすいと感じましたか。
2. 相手と話しやすいと感じましたか。

同室感覚の醸成度

1. 相手を身近に感じることはできましたか。
2. 相手と同じ場所にいるような感じがしましたか。

総合的利用意向

1. このシステムのビジュアル表現が楽しいですか
2. 観光地に設置されていたら利用したいと思いますか

各質問は、“全然そう思わない”から“すごくそう思う”を想定して0~4までの5段階で評価することを被験者に指定して評点を得た。各設問は、単独の質問では、質問文の適切さや意図解釈の揺らぎなどによる誤差が発生するため、各評価項目に関する多面的な質問による総合

評価とすべきであるが、アンケートへの回答の容易性も考慮し、2 個の質問に絞って、揺らぎを抑える効果を期待して、その和をもって評価尺度と定義して分析を行った。

グループ間の比較方法としては、t 検定や F 検定によるパラメトリック検定や多重比較が用いられるが、今回の実験は標本数が少ないことから、母集団分布に正規性・等分散性を仮定することが適切ではないと判断し、ノンパラメトリック検定を用いた。ノンパラメトリック検定の中、比較対象数が 2 群、各群のデータに対応関係がある場合よく用いられるのはウィルコクソンの符号付順位和検定である。本研究は水準間の差の有無を検定するため、有意水準 5% でウィルコクソンの符号付順位和検定を行った。

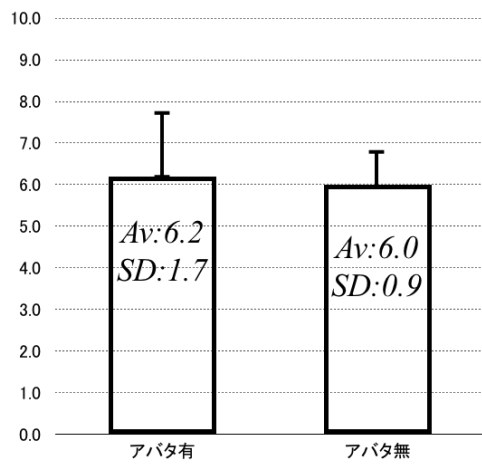
5.5 実験の結果

5.5.1 アバタの効果検証

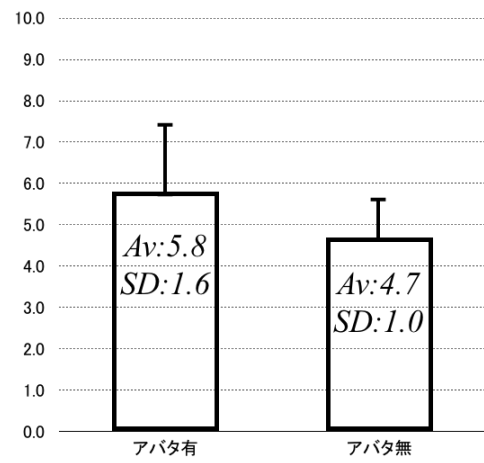
ユーザアバタを用いることで遠隔コミュニケーションに対する影響を検証するために、追加実験のシステム 1 とシステム 2 の実験結果を比べた。タスクの完成時間を表 5.4 に、アンケートの結果を図 5.7 に示す。図 5.7 では各項目の平均値、標準偏差、及び有意差の有無を示す。

タスク達成時間	第 1 グループ	第 2 グループ	第 3 グループ
システム 1	52[s]	22[s]	30[s]
システム 2	73[s]	63[s]	53[s]

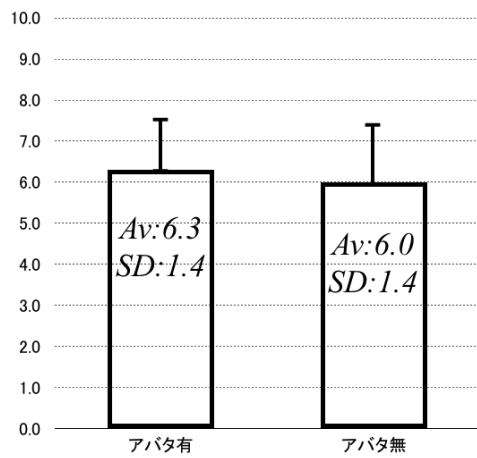
表 5.4 タスクの達成時間



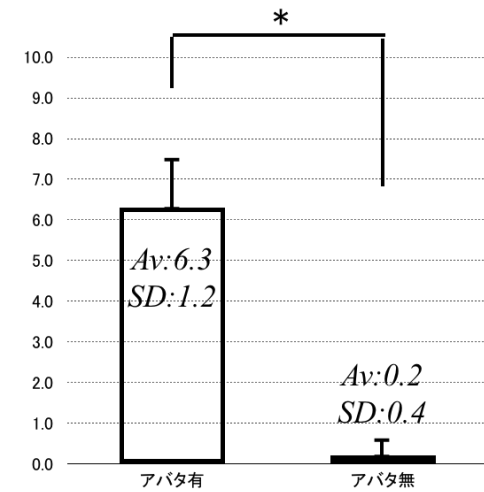
(a) 臨場感



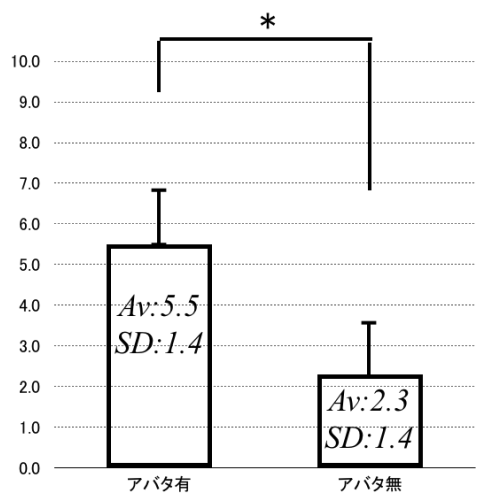
(b) 自由度



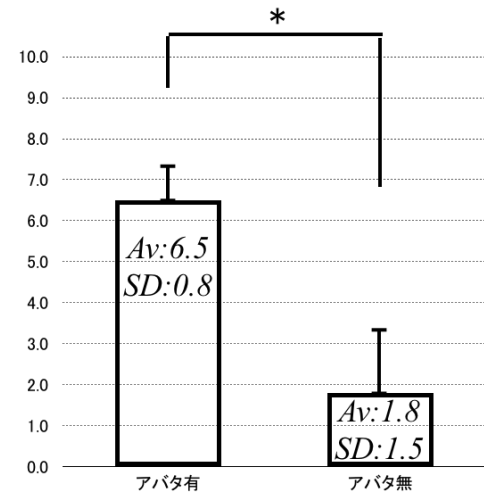
(c) 操作感



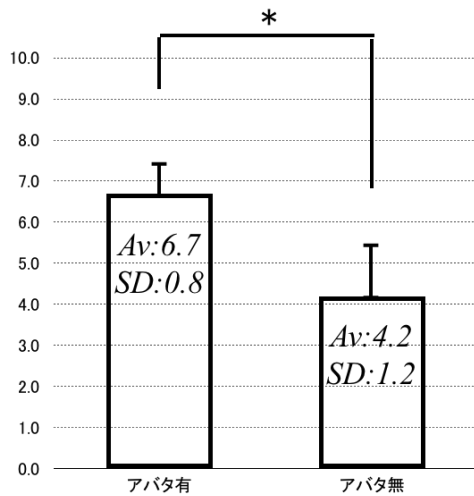
(d) 視線理解容易性



(e) 対話容易性



(f) 同室感



(g) 総合好感度

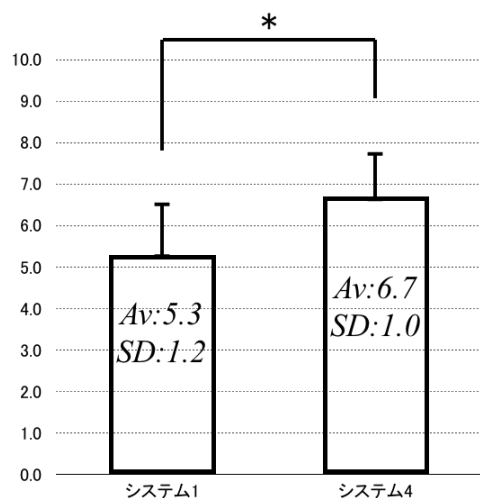
図 5.7 アバタ効果検証実験のアンケートの結果

アバタを用いたシステムは、視線理解容易性、対話容易性、同室感のコミュニケーションのしやすさに関連する3つの項目に、アバタ無しのシステムと有意差が見られた。つまりアバタの使用は、円滑な遠隔コミュニケーションを果たすには機能する。提案手法は有効だと証明された。またアバタを用いたシステムはアバタ無しのシステムと比べ、総合好感度も高い傾向が見られた。実験の様子から、アバタ無しのシステムにおいて、同行者双方は必要な時に言葉を交わす以外、お互いに関心はなかった。それに対しアバタ有りのシステムを利用した時、相手のアバタを見たらそれについていくや、挨拶を交わすなど、コミュニケーションをとる頻度が高かった。

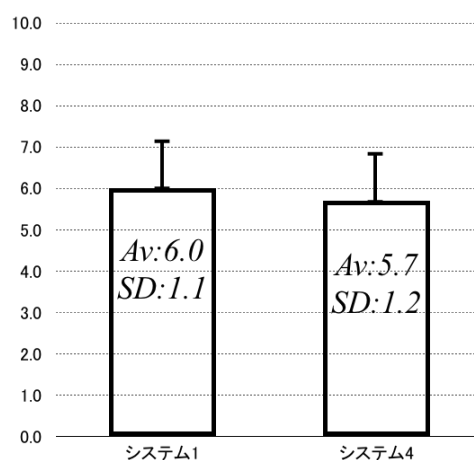
タスク達成時間から、アバタを用いたシステムのほうの時間が比較的短い結果があった。アバタ無しの実験では、被験者の片方が目標を見つけて相手に場所を口で伝えたとしても、相手が場所を特定するまで時間がかかった。対してアバタ有りの実験で目標を見つけっていない被験者は、「見つけた」合図を聞くと、迷わずに相手のアバタに向かって移動し、同じ方向を見ていた。

5.5.2 環境情報更新の効果検証

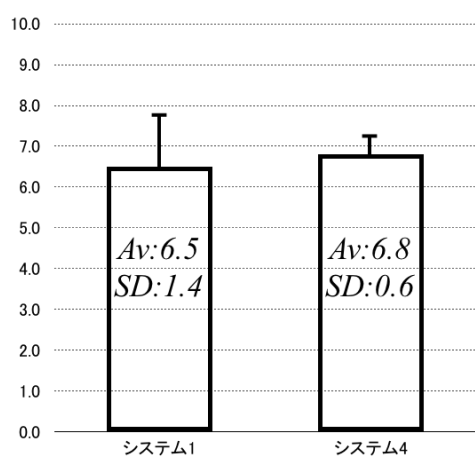
環境情報の更新が遠隔体験及び遠隔コミュニケーションに影響するかどうかを検証するため、1回目のシステム1とシステム4の実験結果を比べた。各項目の平均値、標準偏差、及び有意差の有無を図5.8に示す。



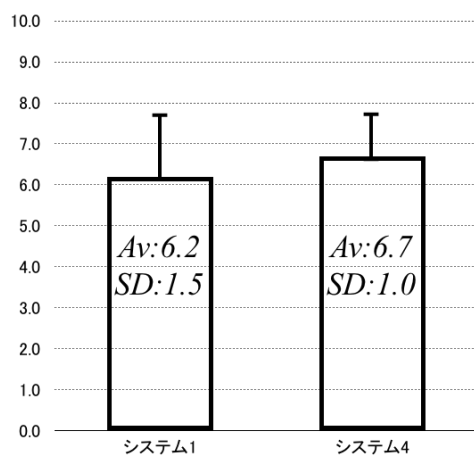
(a) 臨場感



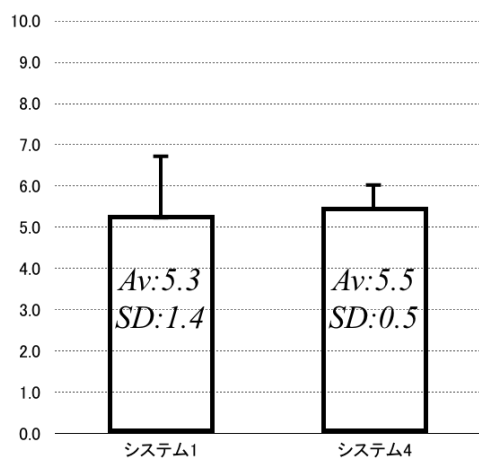
(b) 自由度



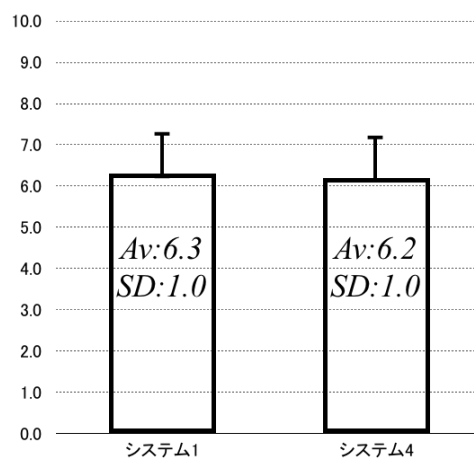
(c) 操作感



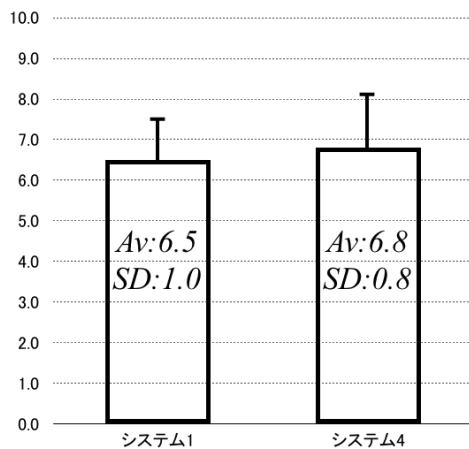
(d) 視線理解容易性



(e) 対話容易性



(f) 同室感



(g) 総合好感度

図 5.8 システム 1 とシステム 4 のアンケートの結果

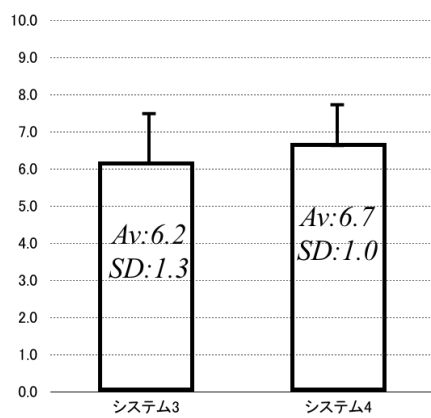
システム 1 とシステム 4 を比べ、環境情報更新の有無の影響を検討するところ、臨場感の評価項目で有意差があった。これは実験場所にあるモニタの動きや、歩行者が歩いていくのを見て差を感じたと考えられる。差は小さいのは、選んだ実験場所は普段あまり人が通り過ぎない場所であるため、映像の変化は少なかったことを原因として思われる。情報更新以外に完全に同じシステムとなっているため、他の項目については特に差がなかった。

5.5.3 自由視点の効果検証

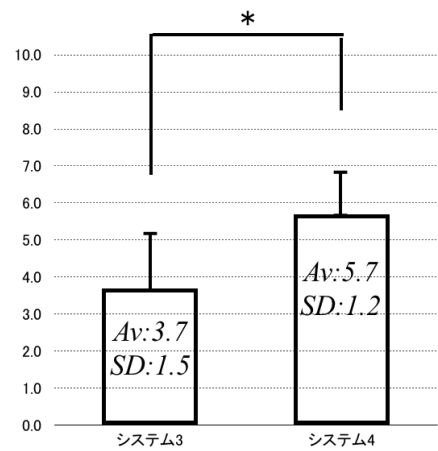
ユーザ視点を変えない場合、操作感にもたらす影響、また第 1 人称視点と第 3 人称視点が見線理解への影響を検証するため、システム 3 とシステム 4 の実験結果を比べた。タスクの達成時間を表 5.5 に、アンケートの結果を図 5.9 に示す

タスク達成時間	第 1 グループ	第 2 グループ	第 3 グループ
システム 3	241[s]	318[s]	205[s]
システム 4	40[s]	43[s]	48[s]

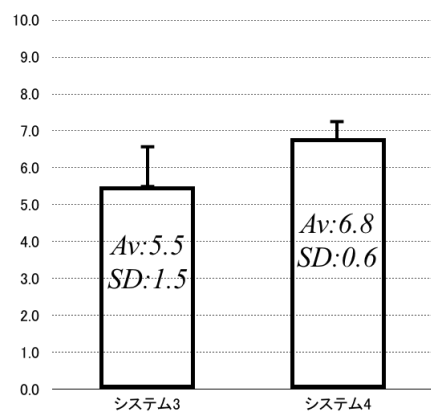
表 5.5 タスクの達成時間



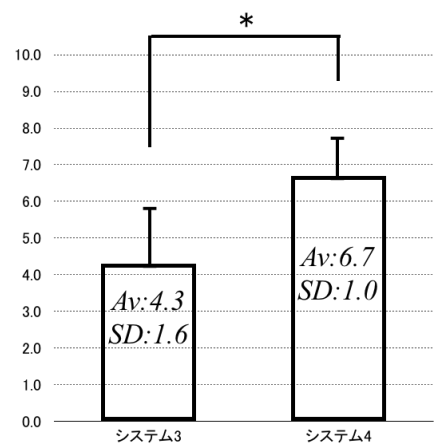
(a) 臨場感



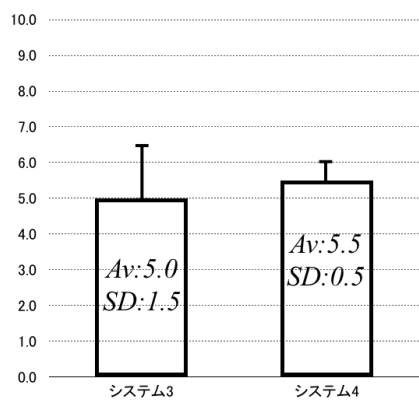
(b) 自由度



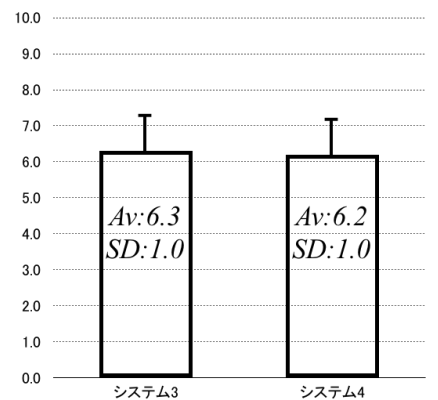
(c) 操作感



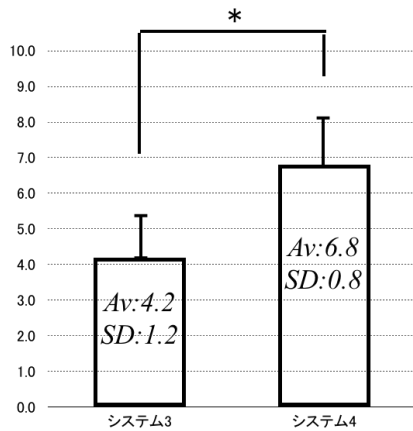
(d) 視線理解容易性



(e) 対話容易性



(f) 同室感



(g) 総合好感度

図 5.9 システム 3 とシステム 4 のアンケートの結果

結果、自由視点なしのほうは色紙を見つけるまでの時間が長かった。

原因として考えるのは、ロボットの操作者はユーザたちのアバタを観察し、操作を行う必要があり、ユーザたちもまたお互いに合わせながら、ロボットの視点範囲内に制限されて自由に動けない。つまり、両方が身動き取りづらい状況となっている。

さらに、このシステムは2人だけではなく、より多くのユーザが同時に使うこととなると、ロボットはアバタに合わせることが非常に難しくなって、ユーザの視点変更はかなり不自由な状態になってしまう。

アンケートによる結果により、自由度と視線理解、総合好感度の項目で有意差があった。臨場感については、どのシステムも最新の情報を提示するため差は出なかった。同室感や対話性について差はないのは、第一人称視点と第三人称視点の違いは、これらの項目に影響が小さいと考えられる。

また、被験者から実験条件毎に得たコメントから、次のような知見が得られた。

臨場感については、第3人称視点の方があると感じたコメントがあった。これは情報の更新ではなく、ロボットが移動していることから映像に激しい変化をもたらし、現場と連動している感覚を被験者に与えたのではないかと考えている。しかし第3人称視点のアンケートに、何人が自由に視点を変えないため不便と感じたというコメントを出した。総合好感度も主にこのことが原因で低くなっていると考えられる。また設定した項目と関係ないコメントで、知人がアバタとなって、リアルのシーンで一緒にゲームをしている感覚がありとても楽しいとのことであった。

第6章 結論

本研究は共同遠隔体験ためのシステムを提案した。いままでのシステムでは同行者を考慮していない点、または情報を更新しないなどの問題点があった。この問題点を改善するため、本研究の仕事は主に以下の通りである。複数人が共同体験する際、円滑にコミュニケーションを取るために環境情報と同行者たちを1つの共同体験バーチャル空間にまとめた。遠隔地にいるロボットが取得した情報に基づき共同体験バーチャル空間を更新し、その映像を自由視点でユーザたちに提示した。

最後にこのシステムを実装し、提案システムと環境情報を更新しないシステム、または提案システムと自由視点がないシステムと比較実験を行った。結果、環境情報を更新するシステムはしないシステムより臨場感はあった。第一人称視点の使用（自由視点あり）は遠隔コミュニケーションに対して視線理解に役立った。また自由視点があるシステムはユーザたちにある程度自由度を与え、いい印象を残った。また実験の様子から、本研究で提案したシステムは環境に明白な変化がある場合だけ、臨場感を与えることができ、そうではない場合、古い画像から作り上げた空間と区別が少なかった。つまり、臨場感を提示するために、動的な環境情報が必要と考えられる。正確さを考量しない場合、ユーザは遠隔地とインタラクションがない以上、昔の映像を利用しても臨場感を与えることはできるのではないかと、まだ検証する必要があるかもしれない。

さらに、アバタの使用は遠隔コミュニケーションに対する影響を検証するため、アバタ有りのシステムと、アバタ無しのシステムと比較実験を行った。アンケートによる結果、コミュニケーションに関連する項目に有意差があった。つまりアバタの使用は本研究の目的の1つである円滑なコミュニケーションに対して有効ということが証明された。

一方、本研究の環境モデルは2次元地図から作られ、環境が複雑になるほど、自由視点画面の矛盾が大きくなってしまう。また大きいシーンにおいて、ロボットのカメラに写す部分だけ環境情報が更新されるため、共同体験バーチャル空間の情報が古すぎないようにするには、ロボットは遠隔地の再築したところを万遍なく巡回するしかない。

謝辞

本研究を進めるにあたり，指導教員の三河正彦准教授には大学院研究生のときから多くのご指導や助言を賜りました．ここまで支えていただいたことに心より御礼申し上げます．また，研究について貴重なご意見を賜りました藤澤誠助教授に感謝致します．三河研究室のOGで私のチューターを担当し，日本に不慣れの時からいろいろ手助けをしてくれた吉川さん，そして，ゼミや研究室の場で多くのご指摘をいただいた三河研究室，藤澤研究室の皆様，実験の際には度々ご協力していただきました．ここに感謝の意を表します．

そして最後に，ここまで私を支え応援していただいた家族に感謝いたします．

曹暢

2016年 3月

参考文献

- [1] Google Map ストリートビュー, <https://www.google.co.jp/intl/ja/maps/streetview/>.
- [2] AFTER DARK, <http://www.afterdark.io/>.
- [3] Oculus Rift, <https://www.oculus.com/>.
- [4] Sanford Dickert, David Maldow, Esg. Robotic Telepresence. Telepresence Options, Summer, pp.50-56, 2013.
- [5] EarthCam, <http://www.earthcam.com/>.
- [6] Knapp, Mark, Judith Hall, and Terrence Horgan. Nonverbal communication in human interaction. Cengage Learning, 2013.
- [7] Billinghurst, Mark, Edward Miller, and Suzanne Weghorst. Collaboration with wearable computers. Augmented Reality, pp.539-577, 2001.
- [8] 石井裕. 身体的アバタを介したコミュニケーション特性の検討 (言語・非言語コミュニケーション~ 言葉と身体との相互作用~). 電子情報通信学会技術研究報告, HCS, ヒューマンコミュニケーション基礎, 108.187 : pp.1-6, 2008.
- [9] 磯和之, 八木貴史, 小林稔, 岩城敏, 石橋聡. 生活融合通信:空間情報整合化機能 “comadapter”. 日本バーチャルリアリティ学会論文誌, Vol. 9 : pp.169-178, 2004.
- [10] Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura, Toru Ishida. FreeWalk: Supporting Casual Meetings in a Network. Proceedings of the 1996 ACM conference on Computer supported cooperative work, pp.308-314, ACM, 1996.
- [11] 仮松田智. アウェアネス情報を取り込んだ仮想空間システムの提案. Unisys 技報, 29.2 : pp.205-216, 2009.
- [12] 有田大作, 吉松寿人, 羽山大介, 谷口倫一郎. Real-time Human Proxy: 仮想空間を用いた遠隔コミュニケーション. 電気学会研究会資料. IP, 情報処理研究会, 2005(1):pp.17-22, 2005.
- [13] 鈴木淳平. Kinect とヘッドマウントディスプレイを用いたバーチャルリアリティ空間共有システム. 大学院研究年報 理工学研究科篇, 42, 2012.
- [14] 柏原忠和, 大澤博隆, 篠沢一彦, 今井倫太. 遠隔コミュニケーションの為の肩乗りアバタの提案. インタラクシオン, 2011 : pp.473-473, 2011.
- [15] 藤村亮太, 郭斌, 大村廉, 今井倫太. 実空間共有型遠隔コミュニケーションシステム Remy の提案. ワークショップ 2008 (GN Workshop 2008) 論文集, Vol. 2008 : pp.85-90, 2008.
- [16] 福島寛之, 鈴木雄介, 細野直恒. 遠隔参加者の存在感を表現する会議ロボットシステム. 人間工学, 44 : pp.300-301, 2008.
- [17] Reina Aramaki, Makoto Murakami. Investigating appropriate spatial relationship between user and ar character agent for communication using AR WoZ system. Proceedings of the 15th ACM on International conference on multimodal interaction : pp.397-404, ACM, 2013.
- [18] 下川裕太, 村上真. 拡張現実を用いた遠隔共食コミュニケーションシステムの構築. 電子情報通信学会技術研究報告, HCS, ヒューマンコミュニケーション基礎, 112.412 : pp.37-42, 2013.
- [19] 三輪敬之, 石引力. 場の創出に影を活用する共存在コミュニケーションシステムの開発. 情報処理学会, インタラクシオン, 2004 : pp.255-262, 2004.

- [20] S. Agarwal, Y. Furukawa, N. Snavely, I. Simon, B. Curless, S. M. Seitz, R. Szeliski. Building rome in a day. Communications of the ACM, Vol.54, No.10 : pp. 105–112, 2011.
- [21] Steven M.Seitz, Charles R.Dyer. View morphing: Uniquely predicting scene appearance from basis images. Image Understanding Workshop, pp.881-887, 1997.
- [22] Marc Levoy, Pat Hanrahan. Light Field Rendering. Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, pp.31-42, ACM, 1996.
- [23] Abe Davis, Marc Levoy, Fredo Durand. Unstructured light fields. Computer Graphics Forum, Vol. 31, No. 2pt1 : pp.305-314, Blackwell Publishing Ltd, 2012.
- [24] Inter Process Communication (IPC), <http://www.cs.cmu.edu/~ipc/>.
- [25] Z. Zhang. A flexible new technique for camera calibration. Trans PAMI, Vol. 22, No. 11 : pp. 1330–1334, 2000.
- [26] 高木幸子, コミュニケーションにおける表情および身体動作の役割. 早稲田大学大学院文学研究科紀要, 第 1 分冊 / 早稲田大学大学院文学研究科編, Vol.5 : pp.25-36, 2005.
- [27] 新徳健, 渡辺富夫. 3 者間インタラクション支援および合成的解析のための身体的バーチャルコミュニケーションシステム. 情報処理学会論文誌, 45.4 : pp.1212-1221, 2004.
- [28] 森田友幸, 間瀬健二, 平野靖, 梶田将司, 岡留剛. ヒューマノイドロボットを用いた遠隔コミュニケーションにおける注目伝達. 情報処理学会論文誌, 48(12):pp.3849-3858, 2007.
- [29] 尾上聡, 山本健太, 田中一品, 中西英之. 遠隔対話者の身体動作の提示による音声コミュニケーションの円滑化. 情報処理学会論文誌, 54(4) : pp.1462-1469, 2013.
- [30] 小木哲朗, 山田俊郎. 複合現実ビデオアバタと仮想空間コミュニケーション. 計測と制御, 41.9 : pp.649-652, 2002.
- [31] MRPT, <http://www.mrpt.org/>.
- [32] OpenCV, <http://opencv.org/>.
- [33] DX ライブラリー, <http://homepage2.nifty.com/natupaji/DxLib/>.
- [34] J. Rollo. Tracking for a Roboceptionist. Undergraduate Senior Research Thesis, Carnegie Mellon University, 2007.
- [35] The Kinect for Windows Software Development Kit, <http://www.microsoft.com/en-us/kinectforwindows/>.
- [36] Segal, Aleksandr, Dirk Haehnel, Sebastian Thrun. Generalized-ICP. Robotics : Science and Systems, Vol. 2, No. 4, 2009.
- [37] ume 式十神白夜, <http://sdrv.ms/MnpedY/>.
- [38] hzeo 式穂歌ソラ, <http://nonbirimattari721.blog.fc2.com/>.
- [39] 細谷英一. 遠隔コミュニケーションにおける自己像を用いた共有空間提示に関する研究. 2011.

付録

実験対象者の方へ

年 月 日

複数ユーザで動の実環境を共有するための 移動ロボットを用いた遠隔体験システム 実験説明書

研究の目的

この研究は、複数人が共に遠隔地を体験する際により楽しい体験ができるよう、お互いコミュニケーションがとりやすく、自由に見渡せ、そして遠隔地の「今」を見ている感じが得られるシステムの構築を目指しています。実験開始前にこの説明書をお読みいただき、ご協力いただける場合には、参加同意書に署名をお願いいたします。

実験方法

画面上に表示された映像を見ながら、自分のアバターを操作し、説明に従ってあるものを探し出していただきます。実験後に、簡単なアンケートをご記入いただきます。所要時間は、前後の説明時間も含めて 1 時間程度です。

個人情報とデータの取扱い

取得したデータや個人情報は、研究目的以外には使用しません。データには番号付けを行うとともに匿名化しますので、専門学会、学術専門誌、学内研究会等を通して研究発表する際も個人情報は守秘されます。データの保管には万全を期し外部へは漏洩しません。この保管データは、研究が終了してから 5 年後までに破棄します。

実験対象者の権利について

この研究に参加するか否かは自由意志で決定してください。また、一度同意した後でいつでも同意を取り消すことができ、それによる不利益はありません。匿名化番号を破棄するとともに、それまでに得られたデータや解析結果を破棄し、それ以降の研究には一切使用いたしません。但し、取り消し要求された時点で公表済みの解析結果がある場合は、このデータを破棄できませんのでご承知おきください。

問い合わせ、苦情等の連絡先

以上，何かご不明な点がありましたら遠慮なくお尋ねください．本研究へのご理解とご協力に深く感謝いたします．

(実験者)

筑波大学大学院 図書館情報メディア研究科
博士前期課程2年 曹 暢 (s1421617@u.tsukuba.ac.jp)

(責任者)

筑波大学 図書館情報メディア系
准教授 三河 正彦 (mikawa@slis.tsukuba.ac.jp)

参加同意書

複数ユーザで動的実環境を共有するための 移動ロボットを用いた遠隔体験システム

<<説明を受けた項目>>

- ☐ 研究の目的
- ☐ 実験方法
- ☐ 個人情報とデータの取扱い
- ☐ 実験対象者の権利について
- ☐ 問い合わせ、苦情等の連絡先
- ☐ 研究代表者の氏名、所属、職名

説明日時： 年 月 日

説明者： _____

責任者: _____

私は、以上の説明を理解し、本研究に参加することに同意します。

年 月 日

所属:

氏 名:

複数ユーザで共同遠隔体験に関する実験のアンケート

システム 1

このシステムのインターフェースは第一人称視点となります。自分のアバターは見えません。画面中にアバターが見えた場合、そのアバターは、この実験であなたと遠隔地を共同体験するもう 1 人の被験者のものです。相手もあなたの視点の位置に、あなたのアバターが見えます。見えるアバターの情報は以下の通りとなります。

- 顔映像
- ジェスチャー（両腕）
- 視点位置
- 視点向き

また、左下は遠隔地のミニマップです。塗り潰した緑の円はあなたの視点の位置、紫色のリングは相手の位置です。

視点の操作方法是以下の通りです。

- ゲームパッド：十字ボタンは視点の回転、右スティックは位置の移動です。
- Kinect：自動的にあなたの両腕のジェスチャーと顔映像をキャッチします。

それでは慣れるまでゲームパッドを使って、自分の視点进行操作し、自由に周りを見渡し、相手とコミュニケーションを取ってください。

システム 4

この実験では、以下の図に示す色紙を探し出してください。壁のどこかに貼っています。2 人ともそれを見つければ、タスク終了となります。先に見つけた方は、タスクが早く達成できるように相手にも場所を教えてください。教える方法は自由です。ゲームパッドの青いボタンを押したらタスクスタートとなります。



色紙の例

システム 3

インターフェースが変わります。この実験は第 3 人称視点となります。自分のアバターも、相手のアバターも一緒に見えます。

操作方法は以下の通りです。

- ゲームパッド：右スティックは位置の移動と回転です。赤いボタンを押したら自分のアバターの位置はカメラ前にリセットされます。
- Kinect：自動的にあなたの両腕のジェスチャーと顔映像をキャッチします。

システム 4 と同じく色紙を探し出し、タスクが早く達成できるよう見つかったら相手にも教えてください。まだ、視点は自分で移動することができません。実験実施者が 2 人のアバターを見て、2 人の移動したい方向を判断し視点を移動させます。自分のアバターが見失ったら四角ボタンを押してからスティックを前に倒しポジションをリセットしてください。

以上で実験は終了です。

アンケートの記入

最後に、ページをめくって、アンケートに答えてください。

追加実験：複数ユーザで共同遠隔体験ためのアバタ効果検証実験のアンケート

システム 1

この実験のインターフェースは第一人称視点となります。自分のアバターは見えません。画面中にアバターが見えた場合、そのアバターは、この実験であなたと遠隔地を共同体験するもう 1 人の被験者のものです。相手もあなたの視点の位置に、あなたのアバターが見えます。見えるアバターの情報は以下の通りとなります。

- 顔映像
- ジェスチャー（両腕）
- 視点位置
- 視点向き

視点の操作方法は以下の通りです。

- ゲームパッド：十字ボタンは視点の回転、右スティックは位置の移動です。
- Kinect：自動的にあなたの両腕のジェスチャーと顔映像をキャッチします。

それでは慣れるまでゲームパッドを使って、自分の視点进行操作し、自由に周りを見渡し、相手とコミュニケーションを取ってください。

続いてタスクを説明します。タスクの内容は、以下の図に示す模様がある場所を探し出すことです。壁のどこかにあります。2 人ともそれを見つけることができれば、タスク終了となります。先に見つけた方は、タスクが早く達成できるように相手にも場所を教えてください、教える方法は自由です。ゲームパッドの青いボタンを押したらタスクスタートとなります。

それでは合図があったら、青いボタンを押して開始してください。



模様の例

システム 2

システム 2 では，システム 1 と同じタスクが設けています．ただし，今回はあなたと共同体験をする相手のアバタが見えないです．

それでは合図があったら，赤いボタンを押して開始してください．

以上で実験は終了です．

アンケートの記入

最後に，ページをめくって，アンケートに答えてください．

アンケート

当てはめる番号にマークをつけてください。

臨場感

Q: 遠隔地の今の様子が分かりましたか

すごく	少し	どちらでも	あんまりない	全然ない

Q: 遠隔地を見ている感じがしましたか

すごく	少し	どちらでも	あんまりない	全然ない

自由度

Q: 現場にいるように自由に見渡せましたか

すごく	少し	どちらでも	あんまりない	全然ない

Q: 自分の行動に制限が多いと感じがしましたか

すごく	少し	どちらでも	あんまりない	全然ない

操作感

Q: 目的地への移動は容易にできましたか

すごく	少し	どちらでも	あんまりない	全然ない

Q: 遠隔地を見渡すための操作は容易にできましたか

すごく	少し	どちらでも	あんまりない	全然ない

視線理解容易性

Q: 相手がどこを見ているか分かりましたか

すごく	少し	どちらでも	あんまりない	全然ない

Q: 相手と目が合っていると感じましたか

すごく	少し	どちらでも	あんまりない	全然ない

対話容易性

Q: 相手とやり取りしやすいと感じましたか

すごく	少し	どちらでも	あんまりない	全然ない

Q: 相手と話しやすいと感じましたか

すごく	少し	どちらでも	あんまりない	全然ない

同室感覚の醸成度

Q: 相手を身近に感じることはできましたか

すごく	少し	どちらでも	あんまりない	全然ない

Q: 相手と同じ場所にいるような感じがしましたか

すごく	少し	どちらでも	あんまりない	全然ない

総合的利用意向

Q: このシステムのビジュアル表現が楽しいですか

すごく	少し	どちらでも	あんまりない	全然ない

Q: 観光地に設置されていたら利用したいと思いますか

すごく	少し	どちらでも	あんまりない	全然ない

このシステムについて感想を自由にご記入ください:

以上で調査は終了です。ご協力ありがとうございました。